

SANDIA REPORT

SAND2017-2854

Unlimited Release

Printed March 2017

xLPR Scenario Analysis Report

Aubrey C. Eckert-Gallup, John R. Lewis, Nevin S. Martin, Lauren B. Hund,
Andrew J. Clark, Dusty M. Brooks, & Paul E. Mariner

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia National Laboratories is a multi-mission laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Approved for public release; further dissemination unlimited.



Sandia National Laboratories

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

NOTICE: This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government, nor any agency thereof, nor any of their employees, nor any of their contractors, subcontractors, or their employees, make any warranty, express or implied, or assume any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represent that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government, any agency thereof, or any of their contractors or subcontractors. The views and opinions expressed herein do not necessarily state or reflect those of the United States Government, any agency thereof, or any of their contractors.

Printed in the United States of America. This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from

U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865) 576-8401
Facsimile: (865) 576-5728
E-Mail: reports@osti.gov
Online ordering: <http://www.osti.gov/scitech>

Available to the public from

U.S. Department of Commerce
National Technical Information Service
5301 Shawnee Rd
Alexandria, VA 22312

Telephone: (800) 553-6847
Facsimile: (703) 605-6900
E-Mail: orders@ntis.gov
Online order: <http://www.ntis.gov/search>



xLPR Scenario Analysis Report

Aubrey C. Eckert-Gallup¹, John R. Lewis², Nevin S. Martin², Lauren B. Hund², Andrew J. Clark³, Dusty M. Brooks¹, & Paul E. Mariner⁴

Departments 6233¹, 436², 6231³, and 6224⁴
Sandia National Laboratories
P.O. Box 5800
Albuquerque, New Mexico 87185-0748

Abstract

This report describes the methods, results, and conclusions of the analysis of 11 scenarios defined to exercise various options available in the xLPR (Extremely Low Probability of Rupture) Version 2.0 code. The scope of the scenario analysis is three-fold: (i) exercise the various options and components comprising xLPR v2.0 and defining each scenario; (ii) develop and exercise methods for analyzing and interpreting xLPR v2.0 outputs; and (iii) exercise the various sampling options available in xLPR v2.0. The simulation workflow template developed during the course of this effort helps to form a basis for the application of the xLPR code to problems with similar inputs and probabilistic requirements and address in a systematic manner the three points covered by the scope.

Keywords: xLPR scenario analysis; mean estimation; sampling confidence intervals; sensitivity analysis; epistemic uncertainty analysis; convergence analysis; recommended sampling scheme

ACKNOWLEDGMENTS

The authors would like to acknowledge the contributions and insights provided by the reviewers of this document including Craig Harrington at Electric Power Research Institute (EPRI), Marjorie Erickson and Bruce Bishop at Phoenix Engineering and Associates Inc. (PEAI), Chris Casarez and Markus Burkardt at Dominion Engineering Incorporated (DEI), Remi Dingreville and Scott Sanborn at Sandia National Laboratories, Cedric Sallaberry at Engineering Mechanics Corporation of Columbus (Emc2), and Jay Wallace and Matthew Gordon at the U.S. Nuclear Regulatory Commission (NRC).

Work performed at Sandia National Laboratories (SNL), a multi-mission laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the U.S. Department of Energy's National Nuclear Security Administration under Contract No. DE-AC04-94AL85000. This report is an independent product of the authors and does not necessarily reflect views held by either SNL or the U.S. Department of Energy.

CONTENTS

1	Introduction.....	27
2	Analysis Methods.....	31
2.1	Separation of Epistemic and Aleatory Uncertainty	32
2.2	Mean Estimation and Sampling Confidence Intervals for a QoI	33
2.3	Sensitivity Analysis.....	43
2.4	Convergence Analysis.....	47
2.5	Epistemic Uncertainty Analysis Methods	49
3	Scenario-Specific Results	55
3.1	Scenario 1	58
3.2	Scenario 2	70
3.3	Scenario 3	74
3.4	Scenario 4	104
3.5	Scenario 5	123
3.6	Scenario 6	141
3.7	Scenario 7	159
3.8	Scenario 8	175
3.9	Scenario 9	196
3.10	Scenario 10	215
3.11	Scenario 11	232
4	Additional Studies.....	245
4.1	Aleatory Sample Size Selection	245
4.2	Importance Sampling Quantile and Methods	251
4.3	Sampling Uncertainty Using Latin Hypercube Sampling.....	260
4.4	SRS Dependence on Random Seed Selection	262
4.5	Bounding Low Probability Events	265
5	Scenario Analysis Summary	271
6	References.....	273
7	Distribution	275

This page intentionally left blank.

FIGURES

Figure 1: Simulation workflow used for the analysis of a single xLPR scenario within the xLPR Scenario Analysis effort.	28
Figure 2: Mean probability of occurrence of axial cracks (solid line) and 95% confidence intervals (dashed lines) for Scenario 3, under two different sampling options. The first (green) uses SRS with a sample size of 100/50 (100 epistemic samples and 50 aleatory). The second run also uses SRS but uses an epistemic sample size of 500.	34
Figure 3: A notional sample of a generic QoI of size $n=40$	36
Figure 4: Bootstrap distribution of sample means for the notional example. The blue dashed vertical lines are the upper and lower bounds of the 95% basic bootstrap confidence interval.	37
Figure 5: 95% confidence interval for the mean probability of axial crack constructed using the nonparametric bootstrap methods.	40
Figure 6: 95% confidence intervals accounting for just epistemic uncertainty (blue) and accounting for both epistemic and aleatory uncertainty (red).	42
Figure 7: Mean probability of axial crack for five independent runs of the xLPR code.	48
Figure 8 : 95% prediction interval for the mean probability of axial crack constructed using the five independent realizations plotted in Figure 7.	48
Figure 9: Sampling confidence intervals on the 95 th percentile of the epistemic realizations for 1000/50 SRS with IS.	51
Figure 10: Probability of occurrence of axial crack for each epistemic realization (grey), the mean (red), and the 0.05 (blue), 0.5 (green), and 0.95 (purple) quantiles.	52
Figure 11: Mean probability (red) and 0.95 quantile (purple) for occurrence of axial leak (solid) and crack (dashed).	53
Figure 12: Additions made to the xLPR model required for the Scenario Analysis report calculations (right). Version control changes are shown in red (left).	56
Figure 13: Mean maximum circumferential crack inner half-length (solid line) and 95% confidence intervals (dashed lines) for Scenario 1, Run 1 (green) and Run 2 (orange).	59
Figure 14: Width of 95% confidence interval for the mean maximum circumferential inner half-length for Scenario 1, Run 1 (green) and Run 2 (orange).	60
Figure 15: Width of 95% confidence interval for the mean maximum circumferential crack inner half-length for Scenario 1, Run 2 (green) and Run 4 (orange).	61
Figure 16: Scatter plots of p9001 (Fatigue Growth Threshold SIF Scaling Factor, CKTH) and the absolute crack depth for Scenario 1, Run 3 (left) and Run 5 (right).	63
Figure 17: Width of 95% confidence interval for maximum circumferential crack inner half-length for Scenario 1, Run 3 (green) and Run 5 (orange).	63
Figure 18: Scatterplot for p9001 (Fatigue Growth Threshold SIF Scaling Factor, CKTH) and maximum circumferential crack depth for Scenario 1, Run 6.	64

Figure 19: Mean maximum circumferential crack depth for Scenario 1, Runs 1 through 6.	65
Figure 20: Mean maximum circumferential crack inner half-length for Scenario 1, Runs 1 through 6.	66
Figure 21: Maximum circumferential crack inner half-length for each epistemic realization (grey), the mean (red), and the 5 th (blue), 50 th (green), and 95 th (purple) percentiles for Scenario 1, Run 3.67	
Figure 22: Maximum circumferential crack depth for each epistemic realization (grey), the mean (red), and the 5 th (blue), 50 th (green), and 95 th (purple) percentiles for Scenario 1 Run 3.....	67
Figure 23: Comparison of estimated mean maximum circumferential crack inner half-length for five convergence replicates for Scenario 1, Run 3.....	68
Figure 24: 95% prediction interval over mean maximum circumferential crack inner half-length for five replicates for Scenario 1, Run 3.....	69
Figure 25: Mean probability of circumferential crack occurrence for Scenario 2 (blue) and Scenario 3 (red).....	71
Figure 26: Mean probability of circumferential leak occurrence for Scenario 2 (blue) and Scenario 3 (red).....	71
Figure 27: Mean probability of circumferential rupture occurrence for Scenario 2 (blue) and Scenario 3 (red).	72
Figure 28: Mean total leak rate [m^3/s] for Scenario 2 (blue) and Scenario 3 (red).	72
Figure 29: Mean probability of occurrence of rupture with ISI for Scenario 2 (blue) and Scenario 3 (red).....	73
Figure 30: Mean probability of occurrence of circumferential ruptures (solid line) and 95% confidence intervals (dashed lines) for Scenario 3, Run 1 (green) and Run 2 (orange).	76
Figure 31: Width of 95% confidence interval for probability of occurrence of circumferential ruptures for Scenario 3, Run 1 (green) and Run 2 (orange).	77
Figure 32: Scatter plots for variable p2543 (multiplier on DM1 proportionality constant A) (plotted on a log scale) and the probability of occurrence for circumferential rupture for Scenario 3, Run 1 (left), Run 2 (center) and Run 3 (right).	77
Figure 33: Scatter plots for variable p2543 (multiplier on DM1 proportionality constant A) (plotted on a log scale) and the probability of occurrence for circumferential cracks for Scenario 3, Run 1 (left), Run 2 (center) and Run 3 (right).	78
Figure 34: Scenario 3, Run 3 scatter plots for variables p2543 (left) (plotted on a log scale), p4352 (center), and p2502 (right).....	78
Figure 35: Number of circumferential rupture events occurring for Scenario 3, Run 3 (No IS) and Run 5 (IS with $q = 0.95$).....	80
Figure 36: Mean probability of occurrence of circumferential ruptures (solid line) and 95% confidence intervals (dashed lines) for Scenario 3, Run 2 (green) and Run 4 (orange).	81

Figure 37: Width of 95% confidence interval for probability of occurrence of circumferential rupture for Scenario 3, Run 2 (green) and Run 4 (orange).	82
Figure 38: Mean probability of occurrence of circumferential ruptures (solid line) and 95% confidence intervals (dashed lines) for Scenario 3, Run 3 (green) and Run 5 (orange).	82
Figure 39: Width of 95% confidence interval for probability of occurrence of circumferential ruptures for Scenario 3, Run 3 (green) and Run 5 (orange).	83
Figure 40: Mean probability of occurrence of circumferential ruptures (solid line) and 95% confidence intervals (dashed lines) for Scenario 3, Run 3 (green) and Run 7 (orange).	84
Figure 41: Width of 95% confidence interval for probability of occurrence of circumferential ruptures for Scenario 3, Run 3 (green) and Run 7 (orange).	84
Figure 42: Mean probability of occurrence of circumferential ruptures (solid line) and 95% confidence intervals (dashed lines) for Scenario 3, Run 7 (green) and Run 9 (orange).	85
Figure 43: Width of 95% confidence interval for probability of occurrence of circumferential ruptures for Scenario 3, Run 7 (green) and Run 9 (orange).	86
Figure 44: Mean probability of occurrence of circumferential ruptures (solid line) and 95% confidence intervals (dashed lines) for Scenario 3, Run 7 (green) and Run 9 (orange) up to 15 years with y-axis shown in log scale.	87
Figure 45: Width of 95% confidence interval for probability of occurrence of circumferential ruptures for Scenario 3, Run 7 (green) and Run 9 (orange) up to 15 years with y-axis shown in log scale.	87
Figure 46: Mean probability of occurrence of circumferential cracks (solid line) and 95% confidence intervals (dashed lines) for Scenario 3, Run 7 (green) and Run 9 (orange).	88
Figure 47: Width of 95% confidence interval for probability of occurrence of circumferential cracks for Scenario 3, Run 7 (green) and Run 9 (orange).	88
Figure 48: Mean probability of occurrence of circumferential ruptures (solid line) and 95% confidence intervals (dashed lines) for Scenario 3, Run 7 (green) and Run 10 (orange).	89
Figure 49: Width of 95% confidence interval for probability of occurrence of circumferential ruptures for Scenario 3, Run 7 (green) and Run 10 (orange).	90
Figure 50: Mean probability of occurrence of circumferential ruptures (solid line) and 95% confidence intervals (dashed lines) for Scenario 3, Run 9 (green) and Run 10 (orange).	91
Figure 51: Width of 95% confidence interval for probability of occurrence of circumferential ruptures for Scenario 3, Run 9 (green) and Run 10 (orange).	91
Figure 52: Mean probability of occurrence of circumferential cracks (solid line) and 95% confidence intervals (dashed lines) for Scenario 3, Run 9 (green) and Run 10 (orange).	93
Figure 53: Width of 95% confidence interval for probability of occurrence of circumferential cracks for Scenario 3, Run 9 (green) and Run 10 (orange).	93

Figure 54: Mean probability of occurrence of circumferential cracks (solid line) and 95% confidence intervals (dashed lines) for Scenario 3, Run 7 (green) and Run 10 (orange).	94
Figure 55: Width of 95% confidence interval for probability of occurrence of circumferential cracks for Scenario 3, Run 7 (green) and Run 10 (orange).....	94
Figure 56: Width of 95% confidence interval for probability of occurrence of circumferential cracks for Scenario 3, Run 7 (green) and Run 11 (orange).....	95
Figure 57: Mean probability of occurrence of circumferential cracks for Scenario 3, Runs 1 through 10.	96
Figure 58: Mean probability of occurrence of circumferential leak for Scenario 3, Runs 1 through 10.	97
Figure 59: Mean probability of occurrence of circumferential rupture for Scenario 3, Runs 1 through 10.	97
Figure 60: Probability of occurrence of circumferential crack for each epistemic realization (grey), the mean (red), and the 5 th (blue), 50 th (green), and 95 th (purple) percentiles for Scenario 3, Run 9.99	
Figure 61: Probability of occurrence of circumferential leak for each epistemic realization (grey), the mean (red), and the 5 th (blue), 50 th (green), and 95 th (purple) percentiles for Scenario 3, Run 9.	100
Figure 62: Probability of occurrence of circumferential rupture for each epistemic realization (grey), the mean (red), and the 5 th (blue), 50 th (green), and 95 th (purple) percentiles for Scenario 3, Run 9.	100
Figure 63: Mean (red) and 95 th percentile (purple) for probability of occurrence of circumferential leak (solid) and crack (dashed) for Scenario 3, Run 9.	101
Figure 64: 95% sampling confidence intervals (dashed) on the 95 th percentile of the epistemic realizations (solid) for the 1000 epistemic samples with LHS and IS.	102
Figure 65: Comparison of estimated mean probability of occurrence of circumferential rupture for five convergence replicates for Scenario 3.	103
Figure 66: 95% prediction interval over mean probability of circumferential rupture for five replicates for Scenario 3.	103
Figure 67: Mean probability of occurrence of circumferential rupture (solid line) and 95% confidence intervals (dashed lines) for Scenario 4, Run 1 (green) and Run 2 (orange).	106
Figure 68: Width of 95% confidence interval for probability of occurrence of circumferential rupture for Scenario 4, Run 1 (green) and Run 2 (orange).	106
Figure 69: Mean probability of occurrence of circumferential rupture (solid line) and 95% confidence intervals (dashed lines) for Scenario 4, Run 3.	107
Figure 70: Scatter plots for variable p2543 (multiplier on DM1 proportionality constant A) and the probability of occurrence for circumferential rupture for Scenario 4, Run 1 (left), Run 2 (center), and Run 3 (right).	108

Figure 71: Number of circumferential rupture events occurring for Scenario 4, Run 3 (NO IS) and Run 4 (IS).	110
Figure 72: Mean probability of occurrence of circumferential rupture (solid line) and 95% confidence intervals (dashed lines) for Scenario 4, Run 3 (green) and Run 4 (orange).	111
Figure 73: Width of 95% confidence interval for probability of occurrence of circumferential rupture for Scenario 4, Run 3 (green) and Run 4 (orange).	111
Figure 74: Mean probability of occurrence of circumferential rupture for Scenario 4, Runs 1 through 5.	112
Figure 75: Mean probability of occurrence of circumferential cracks for Scenario 4, Runs 1 through 5.	113
Figure 76: Probability of occurrence of circumferential crack for each epistemic realization (grey), the mean (red), and the 5 th (blue), 50 th (green), and 95 th (purple) percentiles for Scenario 4, Run 4.	114
Figure 77: Probability of occurrence of circumferential rupture for each epistemic realization (grey), the mean (red), and the 5 th (blue), 50 th (green), and 95 th (purple) percentiles for Scenario 4, Run 4.	115
Figure 78: Mean (red) and 95 th percentile (purple) for occurrence of circumferential leak (dashed) and rupture (solid) for Scenario 4, Run 4.	116
Figure 79: Comparison of estimated mean probability of occurrence of circumferential cracks for five convergence replicates for Scenario 4.	117
Figure 80: Comparison of estimated mean probability of occurrence of circumferential rupture for five convergence replicates for Scenario 4.	117
Figure 81: 95% prediction interval over mean probability of circumferential crack for five replicates for Scenario 4.	118
Figure 82: 95% prediction interval over mean probability of circumferential rupture for five replicates for Scenario 4.	118
Figure 83: Mean probability of axial (dashed lines) and circumferential (solid lines) crack occurrence for Scenario 3 (blue lines) and Scenario 4 (red lines).	119
Figure 84: Mean probability of axial (dashed lines) and circumferential (solid lines) leak occurrence for Scenario 3 (blue lines) and Scenario 4 (red lines).	120
Figure 85: Mean probability of axial (dashed lines) and circumferential (solid lines) rupture occurrence for Scenario 3 (blue lines) and Scenario 4 (red lines).	120
Figure 86: Mean maximum axial (dashed lines) and circumferential (solid lines) iHL for Scenario 3 (blue lines) and Scenario 4 (red lines).	121
Figure 87: Mean maximum axial (dashed lines) and circumferential (solid lines) oHL for Scenario 3 (blue lines) and Scenario 4 (red lines).	121

Figure 88: Mean maximum axial (dashed lines) and circumferential (solid lines) crack depth for Scenario 3 (blue lines) and Scenario 4 (red lines).	122
Figure 89: Mean total leak rate for Scenario 3 (blue line) and Scenario 4 (red line).....	122
Figure 90: Mean probability of occurrence of circumferential rupture (solid line) and 95% confidence intervals (dashed lines) for Scenario 5, Run 1 (green) and Run 2 (orange).	124
Figure 91: Width of 95% confidence interval for probability of occurrence of circumferential rupture for Scenario 5, Run 1 (green) and Run 2 (orange).	125
Figure 92: Mean probability of occurrence of circumferential rupture (solid line) and 95% confidence intervals (dashed lines) for Scenario 5, Run 3.	126
Figure 93: Scatter plots for variable p2543 (multiplier on DM1 proportionality constant A) and the probability of occurrence for circumferential rupture for Scenario 5, Run 1 (left), Run 2 (center), and Run 3 (right).	127
Figure 94: Number of circumferential rupture events occurring for Scenario 5, Run 3 (NO IS) and Run 4 (IS).	129
Figure 95: Mean probability of occurrence of circumferential rupture (solid line) and 95% confidence intervals (dashed lines) for Scenario 5, Run 3 (green) and Run 4 (orange).	130
Figure 96: Width of 95% confidence interval for probability of occurrence of axial cracks for Scenario 5, Run 3 (green) and Run 4 (orange).	130
Figure 97: Mean probability of occurrence of circumferential rupture for Scenario 5, Runs 1 through 5.	131
Figure 98: Mean probability of occurrence of circumferential cracks for Scenario 5, Runs 1 through 5.	132
Figure 99: Probability of occurrence of circumferential cracks for each epistemic realization (grey), the mean (red), and the 5 th (blue), 50 th (green), and 95 th (purple) percentiles for Scenario 5, Run 4.	133
Figure 100: Probability of occurrence of circumferential rupture for each epistemic realization (grey), the mean (red), and the 5 th (blue), 50 th (green), and 95 th (purple) percentiles for Scenario 5, Run 4.....	133
Figure 101: Mean (red) and 95 th percentile (purple) for occurrence of circumferential leak (dashed) and rupture (solid) for Scenario 5, Run 4.....	134
Figure 102: Comparison of estimated mean probability of occurrence of circumferential cracks for five convergence replicates for Scenario 5, Run 4.	135
Figure 103: Comparison of estimated mean probability of occurrence of circumferential rupture for five convergence replicates for Scenario 5, Run 4.	135
Figure 104: 95% prediction interval over mean probability of circumferential cracks for five replicates for Scenario 5, Run 4.	136

Figure 105: 95% prediction interval over mean probability of circumferential rupture for five replicates for Scenario 5, Run 4.	136
Figure 106: Mean probability of axial (dashed lines) and circumferential (solid lines) crack occurrence for Scenario 3 (blue lines) and Scenario 5 (red lines).....	137
Figure 107: Mean probability of axial (dashed lines) and circumferential (solid lines) leak occurrence for Scenario 3 (blue lines) and Scenario 5 (red lines).....	138
Figure 108: Mean probability of axial (dashed lines) and circumferential (solid lines) rupture occurrence for Scenario 3 (blue lines) and Scenario 5 (red lines).....	138
Figure 109: Mean probability of axial (dashed lines) and circumferential (solid lines) crack occurrence for Scenario 4 (blue lines) and Scenario 5 (red lines).....	139
Figure 110: Mean probability of axial (dashed lines) and circumferential (solid lines) leak occurrence for Scenario 4 (blue lines) and Scenario 5 (red lines).....	139
Figure 111: Mean probability of axial (dashed lines) and circumferential (solid lines) rupture occurrence for Scenario 4 (blue lines) and Scenario 5 (red lines).....	140
Figure 112: Mean total leak rate for Scenario 4 (blue line) and Scenario 5 (red line).....	140
Figure 113: Mean probability of occurrence of circumferential ruptures (solid lines) and 95% confidence intervals (dashed lines) for Scenario 6, Run 1 (green) and Run 2 (orange).	143
Figure 114: Width of 95% confidence interval for the estimate of mean probability of occurrence of circumferential ruptures for Scenario 6, Run 1 (green) and Run 2 (orange).....	143
Figure 115: Mean probability of occurrence of circumferential ruptures (solid lines) and 95% confidence intervals (dashed lines) for Scenario 6, Run 2 (green) and Run 3 (orange).	144
Figure 116: Width of the 95% confidence interval for probability of occurrence of circumferential ruptures for Scenario 6, Run 2 (green) and Run 3 (orange).	145
Figure 117: Scatter plots for variable p2543 (multiplier on DM1 proportionality constant A) and the probability of occurrence of circumferential ruptures for Scenario 6, Run 1 (left), Run 2 (center), and Run 3 (right).	146
Figure 118: Number of circumferential rupture events occurring for Scenario 6, Run 3 (No IS) and Run 4 (IS).	147
Figure 119: Mean probability of occurrence of circumferential ruptures (solid line) and 95% confidence intervals (dashed lines) for Scenario 6, Run 3 (green) and Run 4 (orange).	148
Figure 120: Width of 95% confidence interval for probability of occurrence of circumferential ruptures for Scenario 6, Run 3 (green) and Run 4 (orange).	148
Figure 121: Mean Probability of occurrence of circumferential rupture (solid line) and 95% confidence intervals (dashed lines) for Scenario 6, Run 4 (green) and Run 5 (orange).	149
Figure 122: Mean probability of occurrence of circumferential rupture for Scenario 6, Runs 1 through 5.....	150

Figure 123: Probability of occurrence of circumferential crack for each epistemic realization (grey), the mean (red), and the 5th (blue), 50th (green), and 95th (purple) percentiles for Scenario 6, Run 4.	152
Figure 124: Probability of occurrence of circumferential leak for each epistemic realization (grey), the mean (red), and the 5th (blue), 50th (green), and 95th (purple) percentiles for Scenario 6, Run 4.	152
Figure 125: Probability of occurrence of circumferential rupture for each epistemic realization (grey), the mean (red), and the 5th (blue), 50th (green), and 95th (purple) percentiles for Scenario 6, Run 4.	153
Figure 126: Mean (red) and 95th percentile (purple) for occurrence of circumferential leak (solid) and crack (dashed) for Scenario 6, Run 4.	154
Figure 127: Mean (red) and 95th percentile (purple) for occurrence of circumferential rupture (solid) and leak (dashed) for Scenario 6, Run 4.	154
Figure 128: Comparison of estimated mean probability of occurrence of circumferential ruptures for five convergence replicates for Scenario 6, Run 4.	155
Figure 129: 95% prediction interval over mean probability of occurrence of circumferential ruptures for five replicates for Scenario 6, Run 4.	156
Figure 130: Mean probability of axial (dashed lines) and circumferential (solid lines) crack occurrence for Scenario 3 (blue lines) and Scenario 6 (red lines).	157
Figure 131: Mean estimate of total leak rate for Scenario 3 (blue) and Scenario 6 (red).	158
Figure 132: Mean estimate of the probability of occurrence of rupture for Scenario 3 (blue) and Scenario 6 (red).	158
Figure 133: Mean probability of occurrence of circumferential ruptures (solid lines) and 95% confidence intervals (dashed lines) for Scenario 7, Run 1 (green) and Run 2 (orange).	161
Figure 134: Width of 95% confidence interval for the estimate of mean probability of occurrence of circumferential ruptures for Scenario 7, Run 1 (green) and Run 2 (orange).	161
Figure 135: Mean probability of occurrence of circumferential ruptures (solid lines) and 95% confidence intervals (dashed lines) for Scenario 7, Run 2 (green) and Run 3 (orange).	162
Figure 136: Width of the 95% confidence interval for probability of occurrence of circumferential ruptures for Scenario 7, Run 2 (green) and Run 3 (orange).	163
Figure 137: Scatter plots for variable p2543 (multiplier on DM1 proportionality constant A) and the probability of occurrence for circumferential ruptures for Scenario 7, Run 1 (left), Run 2 (center), and Run 3 (right).	164
Figure 138: Mean probability of occurrence of circumferential ruptures (solid line) and 95% confidence intervals (dashed lines) for Scenario 7, Run 3 (green) and Run 4 (orange).	165
Figure 139: Mean Probability of occurrence of circumferential ruptures (solid line) and 95% confidence intervals (dashed lines) for Scenario 7, Run 4 (green) and Run 5 (orange).	166

Figure 140: Mean probability of occurrence circumferential ruptures for Scenario 7, Runs 1 through 5.	167
Figure 141: Probability of occurrence of circumferential crack for each epistemic realization (grey), the mean (red), and the 5th (blue), 50th (green), and 95th (purple) percentiles for Scenario 7, Run 4.	168
Figure 142: Probability of occurrence of circumferential leak for each epistemic realization (grey), the mean (red), and the 5th (blue), 50th (green), and 95th (purple) percentiles for Scenario 7, Run 4.	169
Figure 143: Probability of occurrence of circumferential rupture for each epistemic realization (grey), the mean (red), and the 5th (blue), 50th (green), and 95th (purple) percentiles for Scenario 7, Run 4.	169
Figure 144: Mean (red) and 95th percentile (purple) for occurrence of circumferential leak (solid) and crack (dashed) for Scenario 7, Run 4.	170
Figure 145: Mean (red) and 95th percentile (purple) for occurrence of circumferential rupture (solid) and leak (dashed) for Scenario 7, Run 4.	170
Figure 146: Comparison of estimated mean probability of occurrence of circumferential ruptures for five convergence replicates for Scenario 7, Run 4.	171
Figure 147: 95% prediction interval over mean probability of circumferential ruptures for five replicates for Scenario 7, Run 4.	172
Figure 148: Mean probability of axial (dashed lines) and circumferential (solid lines) leaks occurrence for Scenario 3 (blue lines) and Scenario 7 (red lines).	173
Figure 149: Mean probability of axial (dashed lines) and circumferential (solid lines) rupture occurrence for Scenario 3 (blue lines) and Scenario 7 (red lines).	173
Figure 150: Mean probability of axial (dashed lines) and circumferential (solid lines) crack occurrence for Scenario 3 (blue lines) and Scenario 7 (red lines).	174
Figure 151: Mean estimate of the probability of circumferential ruptures (solid lines) and 95% confidence intervals (dashed lines) for Scenario 8, Run 1 (green) and Run 2 (orange).	177
Figure 152: Width of 95% confidence interval for the estimate of mean probability of circumferential ruptures for Scenario 8, Run 1 (green) and 2 (orange).	177
Figure 153: Mean estimate of the probability of circumferential ruptures (solid lines) and 95% confidence intervals (dashed lines) for Scenario 8, Run 2 (green) and Run 3 (orange).	178
Figure 154: Width of the 95% confidence interval for the mean probability of circumferential ruptures for Scenario 8, Run 2 (green) and Run 3 (orange).	179
Figure 155: Scatter plots for variable p2543 (multiplier on DM1 proportionality constant A) and the occurrence of rupture for Scenario 8, Run 1 (left), Run 2 (center), and Run 3 (right).	180
Figure 156: Mean probability of occurrence of circumferential ruptures (solid line) and 95% confidence intervals (dashed lines) for Scenario 8, Run 3 (green) and Run 4 (orange).	181

Figure 157: Mean probability of occurrence of circumferential ruptures (solid line) and 95% confidence intervals (dashed lines) for Scenario 8, Run 4 (green) and Run 5 (orange).	182
Figure 158: Mean probability of occurrence of circumferential cracks (solid line) and 95% confidence intervals (dashed lines) for Scenario 8, Run 4 (green) and Run 5 (orange).	183
Figure 159: Mean probability of occurrence of circumferential leaks (solid line) and 95% confidence intervals (dashed lines) for Scenario 8, Run 4 (green) and Run 5 (orange).	183
Figure 160: Mean estimate of the probability of occurrence of circumferential ruptures for Scenario 8, Runs 1 through 5.	185
Figure 161: Probability of occurrence of circumferential crack for each epistemic realization (grey), the mean (red), and the 5th (blue), 50th (green), and 95th (purple) percentiles for Scenario 8, Run 4.	186
Figure 162: Probability of occurrence of circumferential leak for each epistemic realization (grey), the mean (red), and the 5th (blue), 50th (green), and 95th (purple) percentiles for Scenario 8, Run 4.	187
Figure 163: Probability of occurrence of circumferential rupture for each epistemic realization (grey), the mean (red), and the 5th (blue), 50th (green), and 95th (purple) percentiles for Scenario 8, Run 4.	187
Figure 164: Mean (red) and 95th percentile (purple) for occurrence of circumferential leak (solid) and crack (dashed) for Scenario 8, Run 4.	188
Figure 165: Mean (red) and 95 th percentile (purple) for occurrence of circumferential rupture (solid) and leak (dashed) for Scenario 8, Run 4.	188
Figure 166: Comparison of estimated mean probability of occurrence of circumferential cracks for five convergence replicates for Scenario 8.	190
Figure 167: 95% prediction interval over mean probability of circumferential cracks for five replicates for Scenario 8.	190
Figure 168: Comparison of the estimated maximum circumferential crack inner half-length for five convergence replicates for Scenario 8.	191
Figure 169: 95% prediction interval over mean estimate of maximum circumferential crack inner half-length for five replicates for Scenario 8.	191
Figure 170: Mean Probability of axial (dashed lines) and circumferential (solid lines) crack occurrence for Scenario 3 (blue lines) and Scenario 8 (red lines).	192
Figure 171: Mean Probability of axial (dashed lines) and circumferential (solid lines) leaks for Scenario 3 (blue lines) and Scenario 8 (red lines).	193
Figure 172: Mean Probability of axial (dashed lines) and circumferential (solid lines) occurrence of rupture for Scenario 3 (blue lines) and Scenario 8 (red lines).	193
Figure 173: Mean estimates of axial (dashed lines) and circumferential (solid lines) maximum crack inner half-length for Scenario 3 (blue lines) and Scenario 8 (red lines).	194

Figure 174: Mean estimates of axial (dashed lines) and circumferential (solid lines) maximum crack outer half-length for Scenario 3 (blue lines) and Scenario 8 (red lines).....	194
Figure 175: Mean estimates of axial (dashed lines) and circumferential (solid lines) maximum crack depth for Scenario 3 (blue lines) and Scenario 8 (red lines).....	195
Figure 176: Mean estimates of axial (dashed lines) and circumferential (solid lines) total leak rate for Scenario 3 (blue lines) and Scenario 8 (red lines).....	195
Figure 177: Mean probability of occurrence of circumferential rupture (solid line) and 95% confidence intervals (dashed lines) for Scenario 9, Run 1 (green) and Run 2 (orange).	197
Figure 178: Width of 95% confidence interval for probability of occurrence of circumferential rupture for Scenario 9, Run 1 (green) and Run 2 (orange).	198
Figure 179: Mean probability of occurrence of circumferential rupture (solid line) and 95% confidence intervals (dashed lines) for Scenario 9, Run 3.	199
Figure 180: Scatter plots for variable p2743 (multiplier on DM1 proportionality constant A) and the probability of occurrence for circumferential crack for Scenario 9, Run 1 (left), Run 2 (center), and Run 3 (right).	200
Figure 181: Scatter plots for variable p2743 (multiplier on DM1 proportionality constant A) and the probability of occurrence for circumferential rupture for Scenario 9, Run 1 (left), Run 2 (center), and Run 3 (right).	200
Figure 182: Number of circumferential rupture events occurring for Scenario 9, Run 3 (NO IS) and Run 4 (IS).	202
Figure 183: Mean probability of occurrence of circumferential rupture (solid line) and 95% confidence intervals (dashed lines) for Scenario 9, Run 3 (green) and Run 4 (orange).	203
Figure 184: Width of 95% confidence interval for probability of occurrence of circumferential rupture for Scenario 9, Run 3 (green) and Run 4 (orange).	203
Figure 185: Mean probability of occurrence of circumferential rupture for Scenario 9, Runs 1 through 5.	204
Figure 186: Mean probability of occurrence of circumferential crack for Scenario 9, Runs 1 through 5.	205
Figure 187: Mean probability of occurrence of circumferential leak for Scenario 9, Runs 1 through 5.	205
Figure 188: Probability of occurrence of circumferential cracks for each epistemic realization (grey), the mean (red), and the 5 th (blue), 50 th (green), and 95 th (purple) percentiles for Scenario 9, Run 4.	206
Figure 189: Probability of occurrence of circumferential rupture for each epistemic realization (grey), the mean (red), and the 5 th (blue), 50 th (green), and 95 th (purple) percentiles for Scenario 9, Run 4.	207
Figure 190: Mean (red) and 95th percentile (purple) for occurrence of circumferential leak (dashed) and rupture (solid) for Scenario 9, Run 4.	208

Figure 191: Comparison of estimated mean probability of occurrence of circumferential cracks for five convergence replicates for Scenario 9.	209
Figure 192: Comparison of estimated mean probability of occurrence of circumferential rupture for five convergence replicates for Scenario 9.	209
Figure 193: 95% prediction interval over mean probability of circumferential cracks for five replicates for Scenario 9.	210
Figure 194: 95% prediction interval over mean probability of circumferential rupture for five replicates for Scenario 9.	210
Figure 195: Mean probability of axial (dashed lines) and circumferential (solid lines) crack occurrence for Scenario 3 (blue lines) and Scenario 9 (red lines).	211
Figure 196: Mean probability of axial (dashed lines) and circumferential (solid lines) leak occurrence for Scenario 3 (blue lines) and Scenario 9 (red lines).	212
Figure 197: Mean probability of axial (dashed lines) and circumferential (solid lines) rupture occurrence for Scenario 3 (blue lines) and Scenario 9 (red lines).	212
Figure 198: Maximum axial (dashed lines) and circumferential (solid lines) iHL for Scenario 3 (blue lines) and Scenario 9 (red lines).	213
Figure 199: Maximum axial (dashed lines) and circumferential (solid lines) oHL for Scenario 3 (blue lines) and Scenario 9 (red lines).	213
Figure 200: Maximum axial (dashed lines) and circumferential (solid lines) crack depth for Scenario 3 (blue lines) and Scenario 9 (red lines).	214
Figure 201: Total leak rate for Scenario 3 (blue line) and Scenario 9 (red line).	214
Figure 202: Mean probability of occurrence of circumferential ruptures (solid line) and 95% confidence intervals (dashed lines) for Scenario 10, Run 1 (green) and Run 2 (orange).	216
Figure 203: Mean probability of occurrence of circumferential ruptures (solid line) and 95% confidence intervals (dashed lines) for Scenario 10, Run 3.	217
Figure 204: Scatter plots for variable p2543 (multiplier on DM1 proportionality constant A) (plotted on a log scale) and the probability of occurrence for circumferential rupture for Scenario 10, Run 1 (left), Run 2 (center) and Run 3 (right).	219
Figure 205: Mean probability of occurrence of circumferential ruptures (solid line) and 95% confidence intervals (dashed lines) for Scenario 10, Run 3 (green) and Run 4 (orange).	220
Figure 206: Width of 95% confidence interval for probability of occurrence of circumferential ruptures for Scenario 10, Run 3 (green) and Run 4 (orange).	220
Figure 207: Width of 95% confidence interval for probability of occurrence of circumferential ruptures for Scenario 10, Run 3 (green) and Run 4 (orange) up to 15 years with y-axis shown in log scale.	221
Figure 208: Mean probability of occurrence of circumferential crack ruptures for Scenario 10, Runs 1 through 5.	222

Figure 209: Mean probability of occurrence of circumferential cracks for Scenario 10, Runs 1 through 5.....	222
Figure 210: Probability of occurrence of circumferential rupture for each epistemic realization (grey), the mean (red), and the 5 th (blue), 50 th (green), and 95 th (purple) percentiles for Scenario 10, Run 4.....	223
Figure 211: Probability of occurrence of circumferential cracks for each epistemic realization (grey), the mean (red), and the 5 th (blue), 50 th (green), and 95 th (purple) percentiles for Scenario 10, Run 4.....	224
Figure 212: Mean (red) and 95th percentile (purple) for occurrence of circumferential rupture (solid) and crack (dashed) for Scenario 10, Run 4.	225
Figure 213: Comparison of estimated mean probability of occurrence of circumferential rupture for five convergence replicates for Scenario 10.	226
Figure 214: Comparison of estimated mean probability of occurrence of circumferential cracks for five convergence replicates for Scenario 10.	226
Figure 215: 95% prediction interval over mean probability of occurrence of circumferential ruptures for five replicates for Scenario 10.	227
Figure 216: 95% prediction interval over mean probability of occurrence of circumferential cracks for five replicates for Scenario 10.	227
Figure 217: Mean probability of axial (dashed lines) and circumferential (solid lines) crack occurrence for Scenario 3 (blue lines) and Scenario 10 (red lines).....	228
Figure 218: Mean probability of axial (dashed lines) and circumferential (solid lines) leak occurrence for Scenario 3 (blue lines) and Scenario 10 (red lines).....	229
Figure 219: Mean probability of axial (dashed lines) and circumferential (solid lines) rupture occurrence for Scenario 3 (blue lines) and Scenario 10 (red lines).....	229
Figure 220: Mean probability of axial (dashed lines) and circumferential (solid lines) crack occurrence for Scenario 4 (blue lines) and Scenario 10 (red lines).....	230
Figure 221: Mean probability of axial (dashed lines) and circumferential (solid lines) crack occurrence for Scenario 8 (blue lines) and Scenario 10 (red lines).....	230
Figure 222: Mean probability of occurrence of circumferential cracks (solid line) and 95% confidence intervals (dashed lines) for Scenario 11, Run 1 (green) and Run 2 (orange).	233
Figure 223: Width of 95% confidence interval for probability of occurrence of circumferential cracks for Scenario 11, Run 1 (green) and Run 2 (orange).	234
Figure 224: Mean probability of occurrence of circumferential cracks (solid line) and 99% confidence intervals (dashed lines) for Scenario 11, Run 1 (green) and Run 2 (orange).	234
Figure 225: Width of 99% confidence interval for probability of occurrence of circumferential cracks for Scenario 11, Run 1 (green) and Run 2 (orange).	235

Figure 226: Mean probability of occurrence of circumferential cracks (solid line) and 95% confidence intervals (dashed lines) for Scenario 11, Run 3 (green) and Run 6 (orange).	236
Figure 227: Width of 95% confidence interval for probability of occurrence of circumferential cracks for Scenario 11, Run 3 (green) and Run 6 (orange).	236
Figure 228: Width of 95% confidence interval for probability of occurrence of circumferential cracks for Scenario 11, Run 4 (green) and Run 6 (orange).	237
Figure 229: Mean probability of occurrence of circumferential cracks (solid line) and 95% confidence intervals (dashed lines) for Scenario 11, Run 5 (green) and Run 7 (orange).	238
Figure 230: Width of 95% confidence interval for probability of occurrence of circumferential cracks for Scenario 11, Run 5 (green) and Run 7 (orange).	238
Figure 231: Scatterplots for variable p2522 (fatigue initiation load sequence factor, FLOAD) and the probability of occurrence for circumferential cracks for Scenario 11, Run 3, 4, and 6.	239
Figure 232: Mean probability of occurrence of circumferential cracks (solid line) and 95% confidence intervals (dashed lines) for Scenario 11, Run 6 (green) and Run 7 (orange).	240
Figure 233: Mean probability of occurrence of circumferential cracks for Scenario 11, Runs 1, 2, 3, 6, and 7.	241
Figure 234: Probability of occurrence of circumferential cracks for each epistemic realization (grey), the mean (red), and the 5 th (blue), 50 th (green), and 95 th (purple) percentiles for Scenario 11, Run 6.	242
Figure 235: Comparison of estimated mean probability of occurrence of circumferential cracks for five replicates for Scenario 11 Run 4 (1000/50 LHS IS 0.95).	243
Figure 236: 95% prediction interval over mean probability of circumferential crack for five replicates for Scenario 11 Run 4 (1000/50 LHS IS 0.95).	243
Figure 237: Empirical (blue) and true (black) CDF and PDF over epistemic samples with a small probability of failure; the true estimate assumes no aleatory uncertainty and the empirical estimate is based on $n_a = 50$	246
Figure 238: Probability of circumferential crack using 100/50 sampling (left) versus 100/1000 sampling (right). Increasing the number of aleatory samples provides a more precise estimate of the failure probability for each epistemic sample.	247
Figure 239: 100 independent estimates (gray) of $P(T_L \leq t \mid R = 5)$ (blue) for aleatory sample size 25 (top right), 50 (top left), 100 (bottom left), and 1000 (bottom right).	248
Figure 240: Figure 3 Dual-loop results for 1000 epistemic samples and aleatory sample sizes of 25 (top left), 50 (top right), 100 (bottom left), and 1000 (bottom right). The blue dashed line is the median of these estimates across the epistemic samples and the green and red lines are the 0.95 and 0.05 quantiles.	249

Figure 241: Efficiency gains using IS as a function of the true probability of occurrence and the targeted IS quantile for the (left) normal-uniform mixture importance distribution and (right) the normal-truncated normal mixture distribution. Both the x- and y-axes are on the log scale. A horizontal line is drawn at 1, where there is no difference in efficiency between SRS and IS. The probability of occurrence is simply a function of μ	252
Figure 242: Probability of circumferential rupture on log scale from time 5 years to 30 years across different SRS sampling schemes with 100 epistemic samples. Black lines correspond to the original random seed and blue lines are a different randomly selected random seed.	254
Figure 243: Probability of circumferential rupture on log scale from time 5 years to 30 years across different SRS sampling schemes with different numbers of epistemic samples.	255
Figure 244: Figure 3: Probability of circumferential rupture on log scale from time 5 years to 30 years across different LHS sampling schemes.	255
Figure 245: Bootstrap relative confidence interval width for probability of circumferential rupture on log scale from time 5 years to 30 years across different SRS sampling schemes with 100 epistemic samples.	256
Figure 246: Bootstrap relative confidence interval width for probability of circumferential rupture on log scale from time 5 years to 30 years across different SRS sampling schemes with different numbers of epistemic samples. We omitted results for $n = 500$ as they were similar to the other sample sizes.	257
Figure 247: Bootstrap relative confidence interval width for probability of circumferential rupture on log scale from time 5 years to 30 years, across different LHS sampling schemes. Confidence intervals are calculated treating the data like an SRS and therefore are not accurate, but relative widths between schemes should be comparable.....	257
Figure 248: SRS 1000/50 runs using no IS versus IS targeting the 0.999 quantile. IS seems to improve inferences from years 5-15 where IS seems to give a smoother and more precise estimate of the failure probability.....	258
Figure 249: Comparison of 10 different 100/50 epistemic/aleatory realizations of the probability of circumferential rupture using LHS and SRS sampling.....	261
Figure 250: Estimated standard deviation of the probability of circumferential rupture across the 10 different realizations for the LHS (blue) and SRS (blue) samples.	261
Figure 251: Mean probability of occurrence of circumferential cracks for Scenario 3, Runs 1 through 10.	262
Figure 252: Median of the sampled uniform random variable from 1000 samples as a function of the epistemic seed (black dots). The white dot and dashed interval reflect 95% confidence intervals on the median from sample of size 1000.	263
Figure 253: Mean probability of occurrence of circumferential cracks for Scenario 3 runs of 500 epistemic, 50 aleatory realizations using SRS (blue) and 1000 epistemic, 50 aleatory realizations using SRS (green) and LHS (purple) with varying random seeds showing the dependence of SRS results on the selected random seed.	264

Figure 254: Distribution of means for each case and sampling strategy. The top plot is Case 1 where T_I is aleatory and R is epistemic. The bottom plot is Case 2 where T_I is epistemic and R is aleatory.	267
Figure 255: Number of samples needed to bound a probability of $1e-6$ at varying levels of confidence.	268

TABLES

Table 1: Typical epistemic QoI vectors for an xLPR output with $n_e = 100$	39
Table 2: Defining features for the 11 xLPR scenarios.....	55
Table 3: Summary of the stepwise rank regression analysis results for all outputs at year 60 for Scenario 1, Run 3.	62
Table 4: Summary of the stepwise rank regression analysis results for all outputs at year 60 for Scenario 3, Run 3.	79
Table 5: Summary of the stepwise rank regression analysis results for all outputs at year 60 for Scenario 4, Run 3	109
Table 6: Summary of the stepwise rank regression analysis results for all outputs at year 60 for Scenario 5, Run 3.	128
Table 7: Summary of the stepwise rank regression analysis results for all outputs at year 60 for Scenario 6, Run 3.	146
Table 8: Summary of the stepwise rank regression analysis results for all outputs at year 60 for Scenario 7, Run 3.	164
Table 9: Summary of the stepwise rank regression analysis results for all outputs at year 60 for Scenario 8, Run 3.	180
Table 10: Summary of the stepwise rank regression analysis results for all outputs at year 60 for Scenario 9, Run 3.	201
Table 11: Summary of the stepwise rank regression analysis results for all outputs at year 60 for Scenario 10, Run 3.	218
Table 12: Summary of the stepwise rank regression analysis for Scenario 11, Run 3 for circumferential crack occurrence by year 60.	239
Table 13: Time (yr) of first circumferential rupture for various sampling schemes.....	253

This page intentionally left blank.

NOMENCLATURE

CDF	Cumulative distribution function
CI	Confidence interval
DM1	Direct model 1
H ₂	Hydrogen
ID	Inner diameter
iHL	Inner half-length
IS	Importance Sampling
ISI	In-service Inspection
LHS	Latin hypercube sampling
LRD	Leak rate detection
MSIP	Mechanical stress improvement process®
oHL	Outer half-length
PCC	Partial correlation coefficient
PRCC	Partial rank correlation coefficient
PWSCC	Primary water stress corrosion cracking
p####	Input data set parameter identification, e.g., p2543
QoI	Quantity of interest
RPV	Reactor pressure vessel
SA	Sensitivity analysis
SIF	Stress Intensity Factor
SRC	Standardized regression coefficients
SRRC	Standardized rank regression coefficients
SRS	Simple random sampling
UA	Uncertainty analysis
WRS	Weld residual stress
xLPR	<u>E</u> <u>x</u> <u>t</u> <u>r</u> <u>e</u> <u>m</u> <u>e</u> <u>l</u> <u>y</u> <u>L</u> <u>o</u> <u>w</u> <u>P</u> <u>r</u> <u>o</u> <u>b</u> <u>a</u> <u>b</u> <u>i</u> <u>l</u> <u>i</u> <u>t</u> <u>y</u> of <u>R</u> <u>u</u> <u>p</u> <u>t</u> <u>u</u> <u>r</u> <u>e</u>
Zn	Zinc

This page intentionally left blank.

1 INTRODUCTION

This report describes the methods, results, and conclusions of the analysis of 11 scenarios defined to exercise various options available in the xLPR (Extremely Low Probability of Rupture) Version 2.0 code. The scope of the scenario analysis is three-fold: (i) exercise the various options and components comprising xLPR v2.0 and defining each scenario; (ii) develop and exercise methods for analyzing and interpreting xLPR v2.0 outputs; and (iii) exercise the various sampling options available in xLPR v2.0. A simulation workflow template (Figure 1) was developed during the course of this effort as a structured approach for the application of the xLPR code to problems with similar inputs and probabilistic requirements and address in a systematic manner the three points covered by the scope. The Scenarios Analysis effort serves to validate the inputs selected for each scenario and, where applicable, recommendations are made for future changes to these inputs that might assist in a better characterization of the scenarios of interest.

In order to accomplish the goals of the xLPR Scenario Analysis effort, statistical methods for post-processing xLPR run output data were developed and implemented in the R statistical software (R Core Team, 2016). The R code developed for this work contains all of the tools required to perform analyses using the methods described in Chapter 2. The use of each of the statistical methods available in this R code provides important information about the uncertainty expected in the results of interest. These uncertainty results are used to compare and differentiate various sampling options exercised for each scenario. Results of the statistical analysis of each xLPR scenario serve as a proof of concept for how these statistical post-processing methods can be applied to similar problems related to the application of xLPR in the future.

The set of simulations run for each scenario were selected strategically and evaluated iteratively so that a comprehensive comparison of sampling options could be made. Scenario 3 served as a test and reference scenario for this process and thus has the largest number of individual simulations. The overall xLPR Scenario Analysis workflow for a single scenario is shown in Figure 1. This process will be described in detail for each of the 11 xLPR Scenarios.

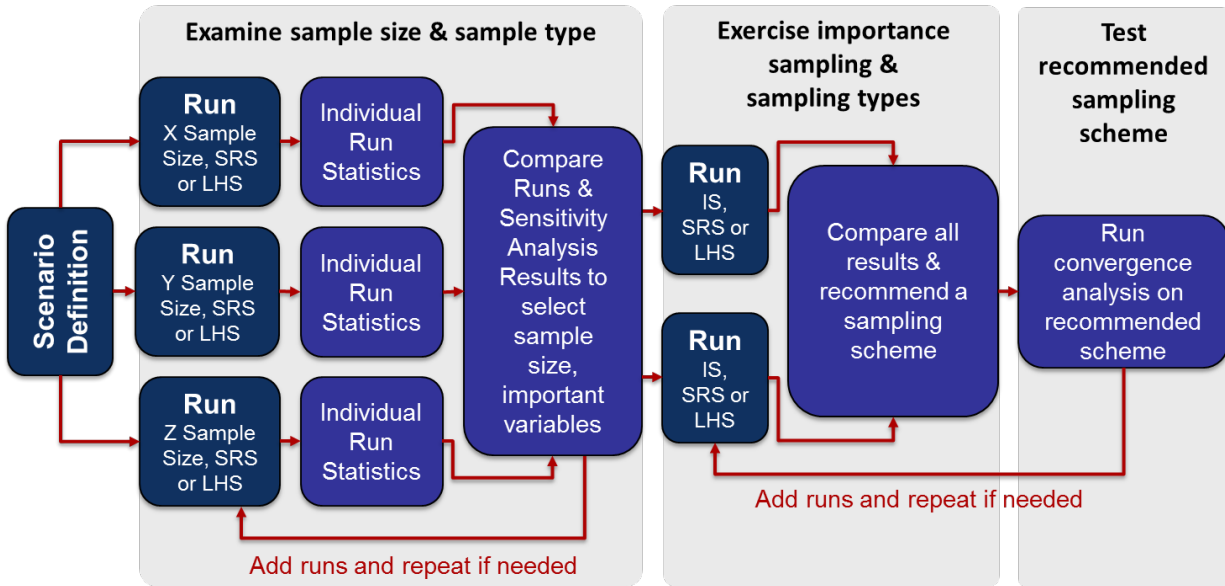


Figure 1: Simulation workflow used for the analysis of a single xLPR scenario within the xLPR Scenario Analysis effort.

The inputs for each scenario were developed by the xLPR Inputs Group. Several additions, described briefly below and in more detail in the introduction of Chapter 3, were made to the beta version of the xLPR V2.0 code in order to generate the data needed to analyze each scenario. These additions were made using the GoldSim version control features and do not impact the calculations or results of the xLPR code. All of the sampled inputs at each realization were exported at the aleatory level (inner sampling loop) so that these sampled inputs could be analyzed along with the outputs of interest using sensitivity analysis techniques. Outputs of interest that were not available in the beta version of the xLPR code, including maximum crack properties at each timestep, were added to the xLPR model interface so that they could be included in the analysis of each scenario. Maximum crack properties are taken as the maximum value (depth, inner half-length, outer half-length) over all subunits at each timestep.

The data and results derived through the use of the simulation workflow template shown in Figure 1 are provided in this report for each scenario. The reduction of epistemic uncertainty, or uncertainty due to a lack of knowledge, was the focus of this work. The overall results of the analysis of all 11 scenarios showed that some sampling selections are more beneficial to the reduction of overall result uncertainty than others. Generally, the following findings were identified:

- Increases in epistemic sample size were found to decrease result uncertainty. The magnitude of this reduction is dependent on the magnitude of the change in epistemic sample size and the output of interest.
- Latin Hypercube Sampling (LHS), as opposed to simple random sampling (SRS), consistently proved to provide more stable coverage of the input space and, therefore, more stable overall results. Because this sampling option does not increase the computational burden required for each simulation and because of its inherent benefits, the use of this sampling option is recommended across all scenarios.

These overall results are supported by the results presented in detail for each scenario in subsequent chapters.

This report is organized as follows. Chapter 2 describes each of the statistical post-processing methods applied throughout the xLPR Scenario Analysis effort and is organized based on the order that each method appears within the analysis of a single scenario. Chapter 3 provides the results of the analysis of each scenario and comparisons amongst scenarios where appropriate. Chapter 4 discusses additional studies with methods that were piloted for a single scenario and broadly referenced or methods that are recommended for similar studies in the future but are beyond the scope of the current work. Chapter 5 provides a concluding summary of the results contained in this report.

This page intentionally left blank.

2 ANALYSIS METHODS

One of the main objectives of xLPR is to reliably estimate failure probabilities under different nuclear power plant environment scenarios (crack initiation and growth mechanisms, mitigation strategies, etc.) (xLPR-SRD-FW, 2016; U.S. NRC, 2015). These failure frequencies are expected to be very small in many applications, which makes it difficult to observe failures from the code in a reasonable amount of computational time. For example, with failure probability $< 1\text{e-}06$ runs, $1\text{e}06$ model runs with randomly sampled inputs will, on average, result in only one observed failure. More sophisticated Monte Carlo sampling techniques, such as importance sampling, can be used to decrease the overall number of required model runs. To construct these more efficient Monte Carlo sampling algorithms, we need to understand which inputs drive variability in the output and are predictive of failure events. While subject matter expertise can be used to identify important inputs, data-driven methods can be applied to determine which inputs drive variability in the output. Given a set of important inputs, specialized sampling techniques like importance sampling (IS) can be applied to improve the efficiency of Monte Carlo resampling (relative to simple random sampling). Once a sampling scheme has been decided upon, statistical analysis methods can be applied to characterize sampling error in the results.

In this section, we describe methods for identifying important model inputs and for quantifying uncertainty associated with the selected sampling scheme. In particular, we describe:

- Separation of aleatory and epistemic uncertainty in the scenario analysis (Section 2.1),
- Estimation and uncertainty quantification for quantities of interest (Section 2.2),
- Sensitivity analysis methods to identify important inputs (Section 2.3),
- Convergence analysis to assess the sufficiency of the selected sampling scheme (Section 2.4), and
- Epistemic uncertainty analysis to quantify the impact of knowledge uncertainty on quantities of interest (Section 2.5).

2.1 Separation of Epistemic and Aleatory Uncertainty

In xLPR, a distinction is made between epistemic and aleatory input variables. Epistemic inputs are uncertain due to knowledge-gaps while aleatory inputs are uncertain due to natural variation. Stated differently, quantification of aleatory uncertainty provides a measure of risk associated with a performance metric, while quantification of epistemic uncertainty provides information on the knowledge about that risk. In the scenario analysis, we separate aleatory and epistemic uncertainties, sampling epistemic uncertainties in an outer loop and aleatory uncertainties in an inner loop. Specifically, we sample n_e unique epistemic inputs in an outer loop. For each sampled epistemic input, we then sample n_a aleatory inputs. The final sample size is then $n_e * n_a$. More information about the treatment of aleatory and epistemic uncertainty can be found in (Helton, Johnson, Sallaberry, & Storlie, 2006) and references therein.

To consider the implications of separating aleatory and epistemic uncertainty, consider estimating a probability, e.g., the probability of circumferential ruptures. Without separating uncertainties, we could run the code n times, calculate the number of ruptures x , and estimate the probability as x/n . If we separate uncertainty, we now have n_a aleatory samples for a single epistemic run. For $i = 1, \dots, n_e$, we can count the number of ruptures x_i in the n_a aleatory samples. So, for epistemic run i , the estimated probability of rupture is x_i/n_a . Separation of uncertainty allows us to assess how these rupture estimates change due to epistemic uncertainty. We can still calculate a ‘best-estimate’ of the probability of rupture over aleatory and epistemic uncertainty, by averaging the n_e different probability estimates.

The advantages of separating aleatory and epistemic uncertainties in the scenario analysis include: (1) when conducting sensitivity analyses, we can marginalize over aleatory uncertainty and restrict to only epistemic variables, reducing the dimension of the inputs (Section 2.3); and (2) we can characterize how much uncertainty in an output is driven by knowledge-gaps (epistemic uncertainty) as opposed to random variation (Section 2.5). Even when uncertainties are separated in running the model, we can average across all epistemic and aleatory samples to calculate a best-estimate (i.e., mean) of the probability of failure as in Sections 2.2 and 2.4.

2.2 Mean Estimation and Sampling Confidence Intervals for a QoI

xLPR outputs multiple quantities for each model run, including binary events such as occurrence of crack, leak, and rupture, as well as continuous measures such as maximum crack length. For each output, we can define a quantity of interest (QoI) summarizing the output over model runs. As an example, for the model output ‘occurrence of circumferential rupture,’ the natural quantity of interest is the probability of circumferential rupture. For continuous measures such as maximum axial crack length, the QoI could be defined as the average max axial crack length, median max axial crack length, or some other summary measure.

The QoI can be calculated for each epistemic sample or can be calculated across all aleatory and epistemic samples. Figure 2 displays examples of a mean estimate and corresponding sampling confidence intervals calculated across all aleatory and epistemic samples for the QoI ‘probability of occurrence of axial crack’. This section describes how such estimates and intervals are constructed. In this section, we focus on the ‘best-estimate’ of the QoI across all aleatory and epistemic samples, along with sampling confidence intervals.

2.2.1 Mean Estimation for a QoI

To obtain a ‘best estimate’ for a QoI, we average over aleatory and epistemic samples. The best-estimate must account for the sampling scheme, as appropriate (e.g., SRS, LHS, importance sampling). Specifically, each observation i has a corresponding weight w_i based on the selected sampling scheme. Under SRS or LHS with no importance sampling, $w_i = 1$ for each i . When importance sampling is applied, the weights correspond to the importance sampling weights.

The specific algorithm for mean estimation is:

1. For each epistemic realization, calculate average the QoI over the n_a aleatory realizations. This quantity, \hat{Q}_i , for $i = 1, \dots, n_e$, is the QoI estimate over the aleatory samples for a specific epistemic realization. (If importance sampling is used for aleatory variables, these weights should be used to calculate the QoI.)
2. Compute the average QoI over the n_e epistemic samples given the QoI estimates \hat{Q}_i and weights w_i for each epistemic sample:

$$\bar{Q} = \frac{\sum_i \hat{Q}_i w_i}{\sum_i w_i}, \quad (1)$$

The estimate \bar{Q} is a ‘best-estimate’ of the QoI over all aleatory and epistemic samples.

2.2.2 Sampling Confidence Intervals for the Mean Estimate

The bootstrap method was used to quantify sampling uncertainty on xLPR QoIs due to the finite number of samples from the model. The bootstrap is a flexible statistical method that allows for the estimation of the sampling distribution of a QoI using resampling methods. As such, it can be used to construct sampling confidence intervals for a QoI. These confidence intervals provide a description of the uncertainty remaining in the true value of the QoI given the finite number of samples available to estimate this QoI.

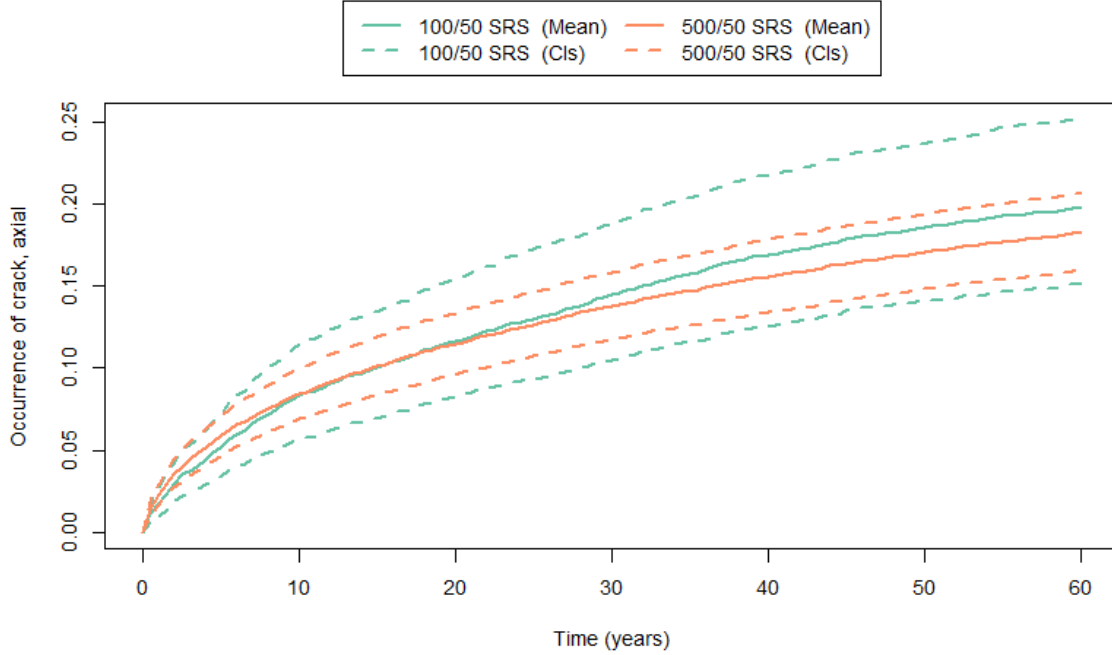


Figure 2: Mean probability of occurrence of axial cracks (solid line) and 95% confidence intervals (dashed lines) for Scenario 3, under two different sampling options. The first (green) uses SRS with a sample size of 100/50 (100 epistemic samples and 50 aleatory). The second run also uses SRS but uses an epistemic sample size of 500.

We implemented a nonparametric bootstrap to quantify sampling uncertainty. More statistical details about the bootstrap method are provided in Section 2.2.2.1. The application of these methods to xLPR outputs is described in Section 2.2.2.2, specifically addressing the functional nature of the outputs (output measured over time) instead of univariate outputs.

2.2.2.1 Nonparametric Statistical Bootstrap Method

In this section, we provide a high-level description of the non-parametric bootstrap (prior to describing the specific bootstrap implementation for xLPR). Consider the objective of estimating the population mean of a generic QoI using a sample of size $n = 40$ pictured in Figure 3. Generally, a point estimate of the population mean as well as a confidence interval (reflecting the uncertainty in the true value of the mean) is desired. A level $1 - \alpha$ statistical confidence interval has the following interpretation: ‘there is $100(1 - \alpha)\%$ confidence that the true value of the parameter falls within the interval.’ Technically, this means that $100(1 - \alpha)\%$ of confidence intervals constructed using repeated sampling from the same population will contain the true parameter, while the other $100\alpha\%$ will not. The value α is referred to as the *significance level* and represents a level of risk that the confidence interval will not contain the true value of the parameter being estimated. Its value is application specific and set with this risk in mind, however, common values are 0.05 and 0.1. The tradeoff with reducing this risk (smaller α) is wider confidence intervals.

There are two general statistical methods for constructing confidence intervals 1) deriving exact or asymptotic approximations to the sampling distribution and 2) approximating the sampling

distribution using resampling methods such as the bootstrap. In the former, an assumption is made about the statistical model from which the sample was generated and deductive properties of a sampling statistic are used to construct the confidence interval for the parameter of interest – in this case the population mean μ . As an illustration, suppose the data (pictured in Figure 3) are sampled independently from a normal distribution with mean μ and variance σ^2 :

$$x_1, \dots, x_n \sim N(\mu, \sigma^2) \quad (2)$$

Under this assumption, the sampling distribution of the sample mean \bar{X} , appropriately centered and scaled, can be shown to follow a Student's t-distribution with $n - 1$ degrees of freedom:

$$T = \frac{\sqrt{n}(\bar{X} - \mu)}{S} \sim t_{n-1}, \quad (3)$$

where S is the sample standard deviation. The above quantity (T) is known as a *pivotal quantity* – a quantity whose sampling distribution does not depend on the unknown model parameters (μ and σ). Since T follows a t-distribution with known degrees of freedom, the following probability statement holds:

$$P\left(\frac{\sqrt{n}(\bar{X} - \mu)}{S} \in \left(t_{n-1, \frac{\alpha}{2}}, t_{n-1, 1-\frac{\alpha}{2}}\right)\right) = 1 - \alpha \quad (4)$$

where $t_{n-1, \gamma}$ is the γ -quantile of the t-distribution and $P(A)$ is the probability of the set A . Rearrangement of the terms inside the probability statement results in a $1 - \alpha$ confidence interval for the population mean μ as simply $\bar{x} \pm t_{n-1, \alpha/2} \frac{s}{\sqrt{n}}$, where \bar{x} and s are the observed sample mean and standard deviation computed from the data. More details on the use of pivotal quantities for constructing confidence intervals can be found in (Casella & Berger, 2002). Given the data in Figure 3, the 95% ($\alpha = .05$) confidence interval for the mean constructed in this way is (9.81, 11.82).

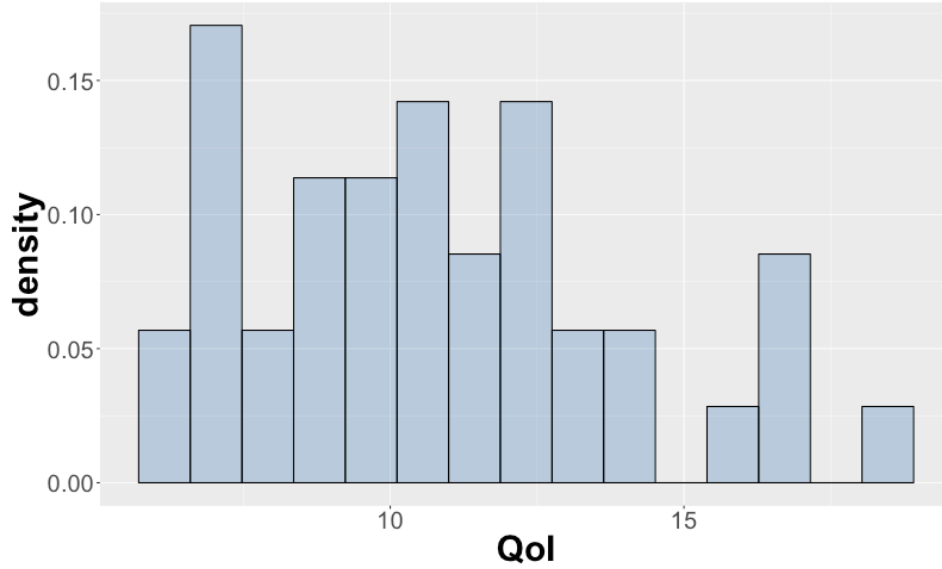


Figure 3: A notional sample of a generic QoI of size $n=40$.

Resampling methods such as the bootstrap provide an alternative sampling based method for constructing confidence intervals that does not require derivation of pivotal quantities for construction of confidence intervals. These bootstrap methods thus have two distinct advantages: ability to calculate uncertainty for arbitrary QoIs without known pivotal quantities, and less reliance on parametric modeling assumptions. There are many different ways to implement the bootstrap to construct confidence intervals, and, herein, we discuss the most frequently used and simplest variety of the bootstrap – the nonparametric bootstrap. For more complete coverage of bootstrap techniques see (Davison, Hinkley, & Young, 2003), (Davison & Hinkley, 1997), and (Efron & Tibshirani, 1994).

The main difference between the nonparametric bootstrap confidence intervals and the pivotal quantity construction of confidence intervals above is that no assumption of the initial sampling distribution is made. Instead, the nonparametric approach estimates the uncertainty of a sampling statistic by resampling the observed data many times, each time computing the statistic of interest, e.g., the mean. In the current notional example, the following steps are repeated many times:

1. Take a random sample, with replacement¹, of size n from the observed data.
2. Compute the sample mean of this random sample.

The sample means from step 2 are collected as pictured in the histogram in Figure 4. The histogram, or bootstrap distribution, is an estimate of the sampling uncertainty of the mean and is analogous to the t-distribution derived in the parametric approach described above. The $100(1-\alpha)\%$ confident *basic* form of the bootstrap interval is $\left(2\bar{x} - q_{1-\frac{\alpha}{2}}, 2\bar{x} - q_{\frac{\alpha}{2}}\right)$ q_{γ} is the γ -quantile of the bootstrap

¹ Sampling ‘with replacement’ means every observed value in the data set is available to be sampled during each of the n sample draws. As a consequence, the same value can appear multiple times in the final sample. This is in contrast to sampling ‘without replacement’ where once sampled, the value is no longer available in subsequent draws.

distribution. For this example, the 95% ($\alpha = 0.05$) confidence interval is (9.83, 11.74) and is shown with vertical dashed blue lines in Figure 4. This interval essentially agrees with the parametric approach. An alternative called the percentile form of the 100(1- α)% confident bootstrap confidence interval is simply $(q_{\frac{\alpha}{2}}, q_{1-\frac{\alpha}{2}})$. The 95% confidence interval constructed in this way is (9.89, 11.80), essentially the same as before due to the symmetry of the histogram in Figure 4. In general, the basic form of the bootstrap confidence interval is preferable to the percentile form of the bootstrap interval.

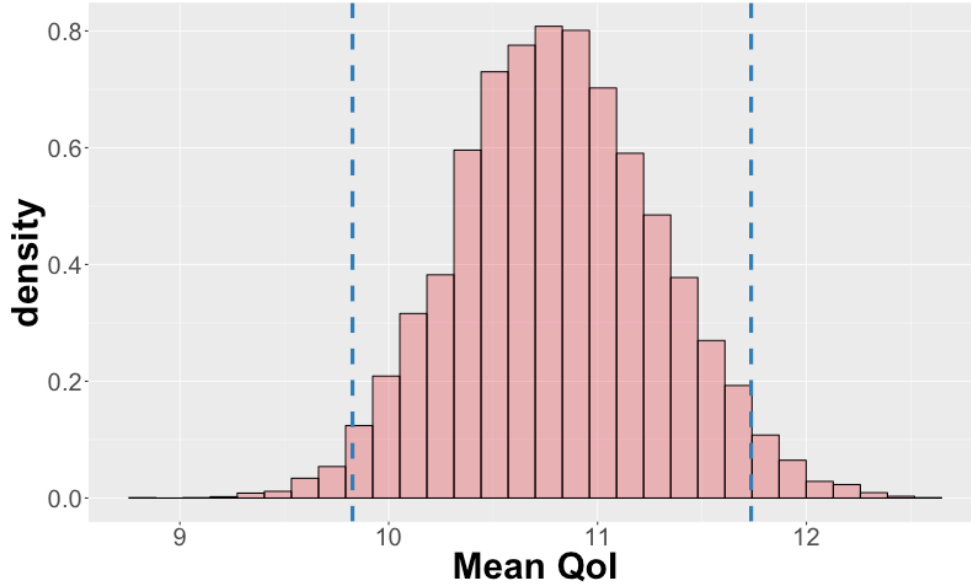


Figure 4: Bootstrap distribution of sample means for the notional example. The blue dashed vertical lines are the upper and lower bounds of the 95% basic bootstrap confidence interval.

This section is intended to describe the traditional nonparametric bootstrap in preparation for the description of the approach used in xLPR. There are many alternative ways to implement a bootstrap, e.g. a parametric bootstrap. For example, if one knows that the data are approximately normally distributed, then making this parametric assumption is warranted and will typically outperform the nonparametric approach. However, exact knowledge of the data generating mechanism is rare and complexity in the data can make parametric assumptions hard to justify. Therefore, in the scenario analysis, we only used the non-parametric bootstrap.

Bootstrap approaches, though flexible, can break down in sparse data situations when there simply is not enough information to estimate the sampling distribution of the data using only the original data. A simple example in the xLPR scenario analysis is, when no events occur, bootstrapping will result in a confidence interval of (0,0). Modifications to the simple nonparametric bootstrap, usually involving additional assumptions (i.e., parametric bootstrap methods), can be made to improve efficiency in some cases. See (Davison & Hinkley, 1997) for more details.

2.2.2.2 Applying the Nonparametric Bootstrap to xLPR QoI

We now describe the bootstrap procedure used to characterize sampling uncertainty in the QoI estimate for xLPR. There are two types of sampling uncertainty: finite aleatory sample size and the finite epistemic sample size. In the scenario analysis, we only account for finite epistemic sample

size in the bootstrap procedure. The bootstrap method accounting for only epistemic sampling uncertainty (outer loop) is described first in Section 2.2.2.3. The bootstrap method that accounts for the finite number of aleatory samples (inner loop) is described in Section 2.2.2.4, but is not implemented in the scenario analysis. This bootstrap requires additional computational resources since one must also save and resample the xLPR output at the aleatory level. In Section 4.1, we discuss the implications of aleatory sampling uncertainty.

2.2.2.3 Bootstrap confidence interval for epistemic uncertainty

To describe the bootstrap method, consider the format of the data displayed in Table 1. For this particular run, the number of epistemic samples, n_e , was set to 100 and the table displays the occurrence of axial crack for each epistemic sample as a function of time, recorded on monthly intervals for the 60-year simulation time. Each column contains results for one epistemic realization and each entry represents the proportion of aleatory samples resulting in a crack at the specific time point in the simulation. The general steps taken in the bootstrap approach are:

1. Take a random sample with replacement of size n_e from the epistemic realizations. If importance sampling was used, the random sample should preserve the data generating mechanism.
2. Use this sample to estimate the desired QoI over time— e.g., compute the sample mean, sample quantiles, etc., at each time point. Note that these are vectors, where the estimates are computed elementwise. If importance sampling is used, the appropriate weights should be applied when estimating the QoI.
3. Iterate steps 1-2 many times, resampling with replacement from the original n_e samples and estimating the QoI for each bootstrap resample. The number of bootstrap resamples is selected based on how much precision is desired in confidence interval estimates, with 1,000 -10,000 iterations typical for confidence interval estimation.
4. Use the collection of estimates over the many iterations to construct the confidence intervals for the desired QoI. Specifically, the basic $(1 - \alpha)100\%$ confident bootstrap interval is $\left(2\hat{\theta} - q_{1-\frac{\alpha}{2}}, 2\hat{\theta} - q_{\frac{\alpha}{2}}\right)$ where $\hat{\theta}$ is the estimate of the QoI using the original data and q_γ is the γ -quantile of the bootstrap distribution, computed elementwise. These are analogous to the dashed blue vertical lines in Figure 4, just computed elementwise across the vectors.

In Step 1, the resampling procedure must be consistent with the way the data were generated. That is, to correctly implement the bootstrap, we need to know how the inputs were sampled and make sure the bootstrap preserves this sampling scheme. For example, the bootstrap cannot be implemented to calculate standard errors for QoIs under Latin hypercube sampling, because it is not possible to preserve the data generating mechanism when re-sampling from the data with replacement. That is, the original inputs were selected using LHS, but we cannot resample from the data using LHS because there is only one observation per LHS stratum. In the scenario analysis, we calculate standard errors for the LHS samples using the SRS bootstrap procedure. Subsequently, these standard errors should not be interpreted as correct estimates of the standard errors under LHS, but rather as standard errors under SRS. These SRS standard errors, though not correct for

LHS, are used to compare relative changes in efficiency according to sample size and importance sampling scheme.

In Step 4, we describe calculation of a two-sided confidence interval, with an upper and lower bound on the probability. For regulatory purposes, one-sided confidence intervals may also be of interest, e.g., an upper bound on the probability of rupture or a lower bound on strength. Such one-sided intervals are trivial to calculate from bootstrap resamples; for instance, a $(1 - \alpha)100\%$ lower confidence bound is $2\hat{\theta} - q_{1-\alpha}$.

Table 1: Typical epistemic QoI vectors for an xLPR output with $n_e = 100$.

<i>month</i>	<i>epistemic sample 1</i>	<i>epistemic sample 2</i>	<i>...</i>	<i>epistemic sample 99</i>	<i>epistemic sample 100</i>
1	0	0	...	0	0
2	0	0	...	0	0
.
.
719	0.36	0.36	...	0.24	0.32
720	0.36	0.36	...	0.24	0.32

As an example, the 95% bootstrap confidence intervals for the mean probability of axial crack using the results of an xLPR run and $n_e = 100$ (with partial epistemic output displayed in Table 1) are displayed in Figure 5. The gray curves in the figure represent the epistemic QoI with pointwise mean represented by the black curve. The bootstrap confidence intervals (accounting for only epistemic sampling uncertainty) are represented by the dashed blue lines.

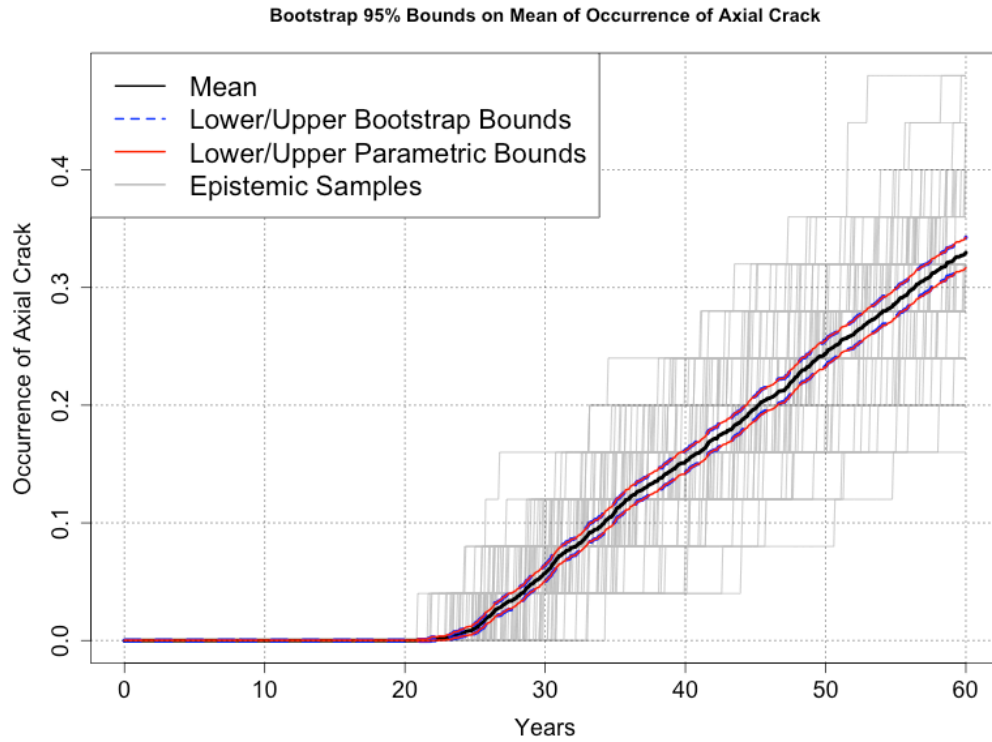


Figure 5: 95% confidence interval for the mean probability of axial crack constructed using the nonparametric bootstrap methods.

2.2.2.4 Bootstrap confidence interval for aleatory and epistemic uncertainty

We now describe the bootstrap procedure used to characterize both epistemic and aleatory sampling uncertainty in the QoI estimate, though this bootstrap is not implemented in the xLPR scenario analysis. To account for aleatory uncertainty in the bootstrap procedure, the output of the xLPR run must be saved at the aleatory level. To explain, each column in Table 1 is the average of the axial

crack output across the n_a aleatory samples. If the individual aleatory output is saved, then aleatory uncertainty in the bootstrap procedure can be included by adding sub-steps to step 1 above:

1. Take a random sample with replacement of size n_e from the epistemic realizations
 - a. For each of the sampled epistemic realizations
 - i. Identify the aleatory realizations that comprise the epistemic realization
 - ii. Bootstrap these aleatory realizations by sampling n_a realizations with replacement and calculating the QoI from the sample. This estimated QoI represents the epistemic realization for the QoI among all the aleatory samples
 - b. Collect the new epistemic sample accounting for aleatory uncertainty found in step 1.a.ii above. Use these in the original step 2 to estimate the desired QoI.

Note that each observed epistemic realization is just an estimate of the QoI (over aleatory samples) given the set of epistemic inputs. The resampling over the aleatory samples is simply a way to estimate the aleatory sampling uncertainty in each epistemic realization. Figure 6 compares the confidence intervals for the mean probability of crack when accounting for just epistemic and when accounting for both epistemic and aleatory uncertainties. As expected, the confidence intervals are slightly larger when accounting for the additional aleatory uncertainty. In this case $n_a = 25$ and the difference between the two bounds will diminish as n_a increases. However, this dependence will be output and QoI dependent. For example, aleatory uncertainty will be more significant if the QoI is an upper quantile (say 0.95) of the output. We discuss the impact of aleatory sampling uncertainty further in Section 4.1.

2.2.2.5 Sampling Confidence Intervals for the Mean Estimate Summary

In this report, we presented a single, simple approach to estimating a bootstrap confidence interval. As mentioned above, there are many other different ways to calculate bootstrap confidence intervals. When examining sampling uncertainty in a QoIs that are probabilities, it is important to interpret confidence interval width relative to the magnitude of the probability. When estimating probabilities, the probability is highly correlated with the variance of the probability. Therefore, confidence interval width is highly correlated with the probability estimate (e.g. probabilities on the order of 10^{-3} will also have confidence interval widths on the order of 10^{-3}). So, when comparing sampling schemes, confidence interval widths should be interpreted relative to the actual estimate.

When bootstrapping, it is important to keep in mind that the procedure requires a sufficiently large sample size to be able to approximate the underlying distribution of the output. This assumption will become particularly important when trying to estimate rare probabilities. For instance, if failure events are observed in only a handful of simulation runs, bootstrapping the failure probability may perform poorly because the underlying distribution of the output is not well-characterized. Additionally, bootstrap confidence intervals cannot be calculated until a failure occurs. When no failure events have happened, no failure events can be resampled from the data, and thus calculated confidence intervals are simply (0,0), which clearly is not correct and should not be interpreted as an actual confidence interval.

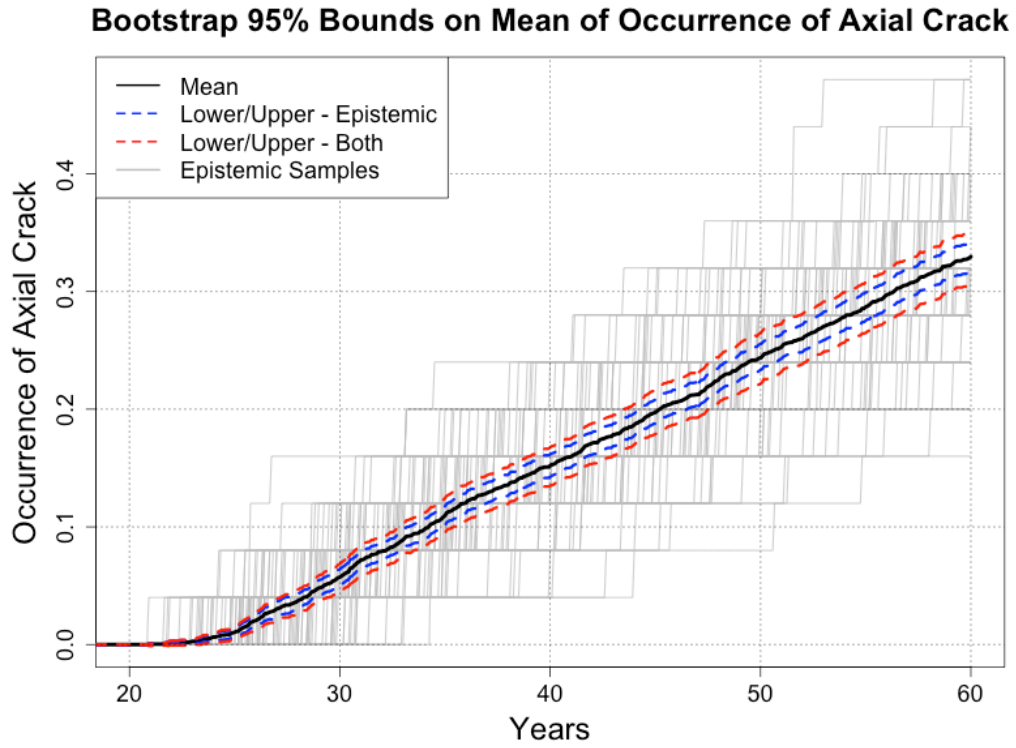


Figure 6: 95% confidence intervals accounting for just epistemic uncertainty (blue) and accounting for both epistemic and aleatory uncertainty (red).

2.3 Sensitivity Analysis

The objective of sensitivity analysis (SA) is to understand the relationship between the model inputs and outputs and, in particular, to identify the inputs that have the most significant impact on the results of the model. The benefits of SA are twofold for xLPR. First, SA helps to increase understanding of the nature of the relationship between the inputs and the outputs. In particular, appropriate SA can quantify the amount of uncertainty in the outputs that is explained by each of the uncertainties in the inputs. Second, the relationship between the most important inputs and the outputs can be examined to determine if there are any regions of the input distribution that can be targeted for importance sampling, which allows for better estimation of small failure probabilities.

For our analysis, we consider SA for epistemic inputs only. We restrict to epistemic variables for two reasons: first, epistemic uncertainties are theoretically reducible and resources could potentially be allocated to reduce the output uncertainty by better understanding these epistemic uncertainties; second, restricting to epistemic parameters reduces the overall number of variables in the SA and also reduces the dimension of the outputs by averaging over the aleatory samples. However, restricting to epistemic variables for the SA has important limitations, including the inability to identify important aleatory variables as well as the inability to identify important interactions between aleatory and epistemic inputs. Specifically, if there is a strong interaction between an epistemic and aleatory variable that results in the epistemic variable having a large impact on the output, such variables may not be detected as sensitive in the analysis. If there are aleatory inputs that are hypothesized to have a large impact on the output, then these inputs can simply be treated as epistemic inputs for the sensitivity analysis.

The SA is performed by assessing the relationship between the sampled inputs and the model outputs; herein, we describe the use of regression techniques for SA. Before discussing regression techniques for SA, the processing and setup of the output for the regression is explained.

To understand the output processing, first recall that two types of uncertain inputs are incorporated in xLPR, 1) epistemic inputs, denoted \mathbf{x}_e and 2) aleatory inputs, denoted \mathbf{x}_a . Uncertain inputs are sampled from their specified probability distribution using a dual-loop algorithm with an inner (aleatory) and outer (epistemic) loop. In the i^{th} iteration of the outer-loop, a vector of epistemic variables $\mathbf{x}_{e,i}$ is sampled from its (joint) distribution. The inner-loop samples n_a (the user-specified aleatory sample size) aleatory variables $\mathbf{x}_{a,1}, \dots, \mathbf{x}_{a,n_a}$ from their probability distribution and runs the code at the joint inputs $(\mathbf{x}_{e,i}, \mathbf{x}_{a,j})$ for $j = 1, 2, \dots, n_a$. The outer loop is then repeated for $i = 1, 2, \dots, n_e$ where n_e is the (user-specified) epistemic sample size. Any output of interest can be generically represented by $f(\mathbf{x}_{e,i}, \mathbf{x}_{a,j})$.

To conduct the SA on epistemic inputs, we first average the output of interest $f(\mathbf{x}_{e,i}, \mathbf{x}_{a,j})$ over the aleatory inputs for each epistemic sample. Letting $y_i = \frac{1}{n_a} \sum_{j=1}^{n_a} f(\mathbf{x}_{e,i}, \mathbf{x}_{a,j})$, the relationship between y_i and the epistemic inputs is modeled statistically by:

$$y_i = f(\mathbf{x}_{e,i}) + \delta_i \quad (5)$$

where $f(\mathbf{x}_{e,i})$ is the model for the expected output (over the distribution of aleatory inputs) of the code and δ_i represents error due to random variations in the observed y_i due to finite aleatory sample size. (Note that the sensitivity analysis could be conducted on the non-averaged response $y_{ij} = f(\mathbf{x}_{e,i}, \mathbf{x}_{a,j})$ instead of y_i if both aleatory and epistemic inputs were included in the sensitivity analysis; we discuss only modeling of y_i herein).

From this point, there are many methods to proceed with the SA. *Sensitivity measures* are metrics which quantify the relationship between the inputs and outputs. Perhaps the simplest approach to calculating a sensitivity measure is to use multiple linear regression. Let $\mathbf{x} = (x_1, x_2, \dots, x_p)$ represent the random vector of all p uncertain epistemic inputs (the fixed inputs will not contribute to the regression analysis and are left out). Let \mathbf{X} be the $n_e \times (p + 1)$ *design matrix* whose i^{th} row is $(1, \mathbf{x}_i)$ with $\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{ip})$ (where x_{ij} is just the i^{th} observed value of the j^{th} input). The multiple linear regression model is:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \quad (6)$$

where $\mathbf{y} = (y_1, \dots, y_{n_e})^T$, $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_p)^T$ is a vector of unknown regression coefficients, and $\boldsymbol{\epsilon}$ are independent, mean-zero residual errors (accounting for finite aleatory sample size and variation in the output that cannot be explained using the linear model). The coefficient β_0 represents the intercept and the coefficient β_1 , for example, represents the expected change in the response given one unit increase in the input x_1 , fixing the remaining inputs (x_2, \dots, x_p) .

The model in Equation (6) assumes that all inputs vary linearly with the output and that there are no interactions in the relationship between the inputs and output. The multiple linear regression model can be extended to include more complicated parametric relationships, e.g., $\mathbf{y} = g(\mathbf{X})\boldsymbol{\beta} + \boldsymbol{\epsilon}$, where $g(\mathbf{X})$ is a transformation of the input variables and $\boldsymbol{\beta}$ is again a vector of regression coefficients; this model formulation can encompass interactions between inputs, polynomial effects, and smoothing splines. Further, there are alternative sensitivity analysis techniques that can accommodate interactions and non-linearities, such as variance-based metrics based on sampling directly from the model or building flexible emulators to approximate the model (Saltelli, Chan, & Scott, 2000) (Storlie, Swiler, Helton, & Sallabery, 2009).

For sake of simplicity and interpretability, we stay within the multiple linear regression framework for the scenario analysis. To relax the linearity assumption, we use *rank regression* to accommodate nonlinear, monotonic input-output relationships in the SA. Rank regression is a statistical regression tool for quantifying the strengths of relationships between inputs and outputs in nonlinear situations as long as these relationships are (approximately) monotonic (Iman & Conover, 1979). Rank regression is simply a regression on the ranks of the data. The form of the multiple linear regression model in Equation (6) is the same, except the response and the inputs are replaced by their corresponding ranks. Ranking refers to sorting numerical data (smallest to largest) and replacing the numerical value with its rank in the ordering. Ties can occur if numerical values are the same. This often happens with ‘occurrence’ responses in xLPR (crack, leak, and rupture) as zero occurrences across all aleatory runs often occurs. In tie situations, the average rank is used. For example if 10 of the observed response values are zero and all others nonzero, their ranks are 1,2,...,10. The average rank of 5 is assigned as the rank for the rank regression.

Given the multiple linear regression model of the form in Equation (6), we now discuss the sensitivity analysis measures that we use for the SA. For multiple regression, the model coefficient estimates are:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{y}. \quad (7)$$

Recall that $\hat{\beta}_j$ is an estimated average rate of change in \mathbf{y} as a function of x_j (fixing the remaining variables) and has units corresponding to the units of \mathbf{y}/x_j . As a result, one would not use $\hat{\boldsymbol{\beta}}$ as currently specified as a sensitivity measure because its elements are dimensioned. That is, the magnitude of $\hat{\beta}_j$ is dependent on the magnitude of x_j and \mathbf{y} . These coefficients are thus normalized so that they can be directly compared:

$$\hat{\beta}_j^* = \hat{\beta}_j \frac{s_{x_j}}{s_y}, \quad (8)$$

where s_z^2 represents the (sample) variance of the variable z . The $\hat{\beta}_j^*$ are referred to as the (estimated) standardized regression coefficients (SRCs). When using rank regression, these coefficients are called standardized rank regression coefficients (SRRCs).

In practice the SRCs (and SRRCs) are estimated as follows. First the inputs and outputs are normalized:

$$y_i^* = \frac{y_i - \bar{y}}{s_y}, \quad x_{ij}^* = \frac{x_{ij} - \bar{x}_j}{s_{x_j}}, \quad (9)$$

where $\bar{y} = \frac{1}{n_e} \sum_{i=1}^{n_e} y_i$, $s_y = \frac{1}{n_e-1} \sum_{i=1}^{n_e} (y_i - \bar{y})^2$, x_{ij} is the i^{th} observation of input j , $\bar{x}_j = \frac{1}{n_e} \sum_{i=1}^{n_e} x_{ij}$, and $s_{x_j} = \frac{1}{n_e-1} \sum_{i=1}^{n_e} (x_{ij} - \bar{x}_j)^2$. Letting $\mathbf{y}^* = (y_1^*, \dots, y_{n_e}^*)^\top$ and \mathbf{X}^* be the $n_e \times p$ design matrix whose i^{th} row is $\mathbf{x}_i^* = (x_{i1}^*, x_{i2}^*, \dots, x_{ip}^*)$, the regression in (6) becomes:

$$\mathbf{y}^* = \mathbf{X}^* \boldsymbol{\beta}^* + \boldsymbol{\epsilon}^* \quad (10)$$

where $\boldsymbol{\beta}^* = (\beta_1^*, \dots, \beta_p^*)^\top$. After standardizing the coefficients, there is no more need for an intercept term (β_0 and the vector of 1's in the first column of \mathbf{X}) since the centering and scaling in (9) implies the intercept is zero. The values of the estimated $\hat{\beta}_j^*$ in (8) are the same as the corresponding element in $\hat{\boldsymbol{\beta}}$ in equation (7) with \mathbf{y} and \mathbf{X} replaced with \mathbf{y}^* and \mathbf{X}^* , respectively.

The SRC represents an estimate of the expected number of standard deviations the output will change per standard deviation change in the input variable (while holding all other input variables constant). In this sense, the (squared) SRCs describe how much each individual factor contributes to the overall variance of the output. As a sensitivity measure, their relative magnitudes indicate their order of importance.

Another useful sensitivity measure is the *partial correlation coefficient* (PCC) which measures the degree of association between the output and a single input, with the effects of the other inputs

removed. For rank regression, these coefficients are called and the *partial rank correlation coefficient* (PRCCs). To compute the PCC (or PRCC) between y and, say, x_1 given x_2, \dots, x_p , one just computes the correlation between the residuals from two linear regressions 1) y with x_2, \dots, x_p , and 2) x_1 with x_2, \dots, x_p . As correlations, PCCs can take on values between -1 and 1 with values closer to the extremes indicating stronger (linear) associations. When the inputs are independent, both SRC and PCC should give the same results in terms of which variables have the most influence on the output.

Prior to estimating sensitivity measures, one must choose the set of input variables to include in the regression analysis. A simple tool for variable selection is stepwise regression. Stepwise regression is a variable selection technique that iteratively drops variables that are not statistically significant and adds statistically significant variables until an overall reduced model is found, where statistical significance is determined based on some specific statistical criteria (Neter, Kutner, Nachtsheim, & Wasserman, 1996). The SRCs or PCCs from the final reduced model are used for sensitivity measures.

If the user aims to identify variables for importance sampling using data-driven methods, the regression methods herein are a simple starting point. Specifically, the user could start with a rank regression model and/or a multiple linear regression model and use the stepwise procedure for variable selection to reduce the dimension of the input parameter space for sensitivity analysis. Simple scatterplots of the output versus the inputs can also be used to identify strong correlations between inputs and outputs. In the case where more refined modeling is need (e.g. to accommodate variable interactions), more general sensitivity metrics measuring the relative variance in the response explained by each input (or a group of inputs) can be computed (Saltelli, Chan, & Scott, 2000). As an example, (Storlie, Swiler, Helton, & Sallabery, 2009) describe a method for computing variance-based metrics using flexible regression models.

2.4 Convergence Analysis

When using the results of outputs from computer models such as xLPR it is important to assess if the code has been run using a large enough sample size under the sampling options to ensure the convergence of the results (i.e. a larger sample size will not improve estimation on the QoI). The assessment of convergence is both QoI and application specific. For example, the uncertainty in the mean of a particular output may be assessed to be sufficiently small for the given application. Hence, the result would be deemed to have converged. However, the uncertainty in the mean of another output, or a quantile of the same output could be different. It is not the intent of this section to provide yes or no criteria on run convergence, since such a decision is application- and QoI-specific. Rather, we describe methods used to quantify run-to-run variability in the mean output in the scenario analysis to assess convergence in the scenario analysis.

One way to assess potential run-to-run variability in a QoI is to construct confidence intervals on the QoI for a single run using the approaches in Section 2.2.2. Such intervals provide a range in which it is expected with $1 - \alpha$ confidence that the true QoI lies. (Recall that these confidence intervals will not be accurate for LHS; however, they will provide an upper bound on sampling uncertainty under LHS, given that LHS should be at least as efficient as SRS.)

In the scenario analysis, we select a sampling option for which the confidence interval is sufficiently narrow. Then, we re-ran the code under this sampling option several independent times and compare the results across the different runs. Re-running the code acts as an independent check of convergence. (In practice, if multiple convergence runs are available, they can be combined to improve precision in QoI estimates.) If the variability of the QoI across the independent runs is sufficiently small, then one could conclude that the result for this QoI has converged. To use this approach, a sufficiently large number of runs must be conducted to be able to characterize variability across runs; herein, we use 5 runs. Ideally, to ensure independence of runs, the random number seeds (both epistemic and aleatory) for each run should be unique and should be randomly generated. In the scenario analysis convergence runs, we did not generate the random seeds in a random fashion, which may have induced correlations between runs, as discussed in Section 4.4.

Computational resources limit the number of independent runs that can be feasibly obtained; statistical prediction intervals can be used to statistically bound where the QoI for an additional run may lie. We construct prediction intervals assuming the QoI is normally distributed, which should be a reasonable assumption when the QoI is a mean across model runs (but will likely perform poorly for rare probabilities). Under this assumption, the prediction interval is:

$$\bar{x} \pm t_{\alpha/2, n_r-1} \sqrt{s^2 \left(1 + \frac{1}{n_r}\right)}, \quad (11)$$

where n_r is the number of independent replicates (5 in this case), and \bar{x} and s are the sample mean and standard deviation of the replicates, respectively (Casella & Berger, 2002). A $1 - \alpha$ *prediction interval* is interpreted as ‘there is $1 - \alpha$ confidence that the QoI from an additional code run will lie between these bounds (assuming the statistical model is correct).’ As an example, Figure 7 displays the mean probability of axial crack across five independent runs of the xLPR code. A prediction interval using the five independent runs is in Figure 8. The width of this interval gives a

quantitative metric of where it is expected the QoI of a future independent run will lie. That is, convergence of the QoI is assessed based on the prediction interval width.

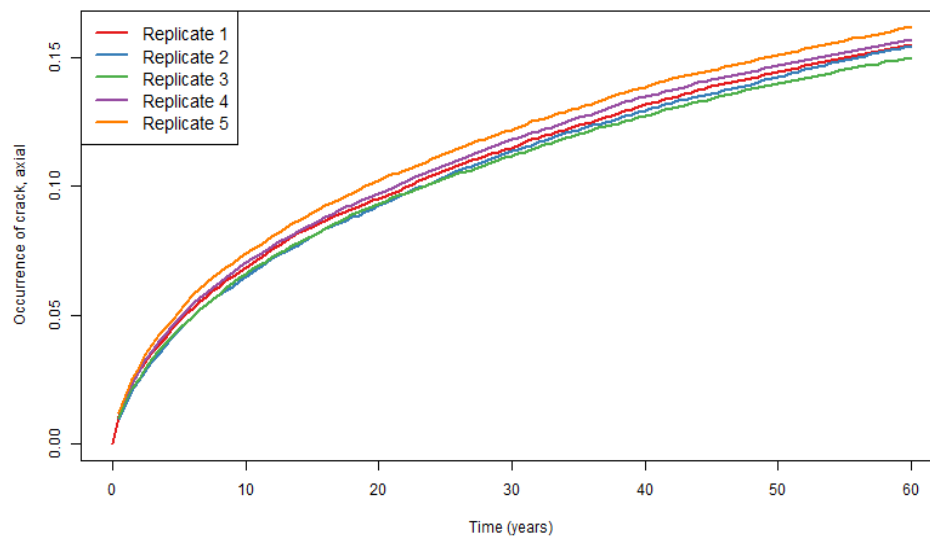


Figure 7: Mean probability of axial crack for five independent runs of the xLPR code.

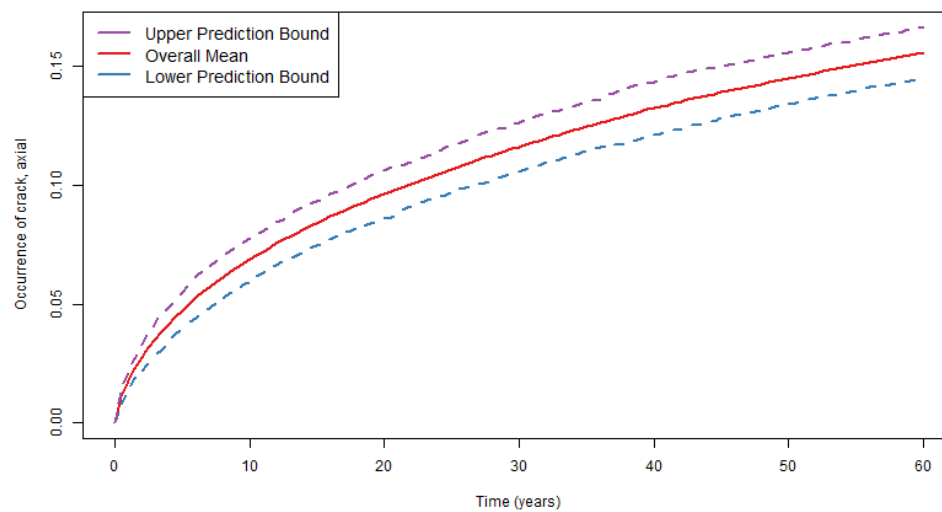


Figure 8 : 95% prediction interval for the mean probability of axial crack constructed using the five independent realizations plotted in Figure 7.

2.5 Epistemic Uncertainty Analysis Methods

The purpose of epistemic uncertainty analyses (UA) is to understand the uncertainty in the QoI estimates across epistemic samples in order to quantify knowledge uncertainty in the QoIs (risk estimates). In general, UA consists of analyzing the outputs of interest over epistemic samples from a sufficiently converged model run. UA results can be quantified by calculating quantiles of the QoI over the epistemic samples.

2.5.1 Epistemic Quantiles

To compute epistemic quantiles, we simply use empirical quantile estimates over the epistemic runs. The specific algorithm for quantile estimation is:

1. For each epistemic realization, calculate the QoI over the n_a aleatory realizations. This quantity, \hat{Q}_j , for $j = 1, \dots, n_e$, is the QoI estimate over the aleatory samples for a specific epistemic realization.
2. Compute the desired quantile of the QoI over the n_e epistemic samples.

2.5.2 Quantile Estimation

When importance sampling is not used, calculating empirical quantiles is straightforward using observed order statistics. For instance, when $n_e=100$, the 5th and 95th quantile estimates are simply the 5th and 95th largest values of the QoI. When importance sampling is used, the importance sampling weights must be applied in the analysis. We describe the approach we used for quantile estimation with a weighted sample. Many variants of quantile estimates exist (Hyndman & Fan, 1996) but many take the form of a weighted average of two order statistics:

$$q = (1 - \gamma)y_{(l)} + \gamma y_{(h)}. \quad (12)$$

In the above formula, γ is a weight with value between 0 and 1, and subscripts in parentheses indicate order statistics. For example, $y_{(1)}$ is the smallest value in the sample, $y_{(2)}$ the second smallest, and so on.

For the scenario analysis, we use the default weighted quantile estimator in the R package Hmisc (Harrell, 2016). The technical details for this estimate are described below. Let $S_k = \sum_{i=1}^k w_{(i)}$ for $k = 1, 2, \dots, n_e$ where $w_{(i)}$ is the weight of the i^{th} order statistic $y_{(i)}$. The S_k are simply the cumulative sums of the ordered weights. Next set $\eta = 1 + (S_{n_e} - 1)q$ and use the weighted average formula above with:

$$\begin{aligned} l &= \operatorname{argmin}_k S_k \geq l', l' = \max(\lfloor \eta \rfloor, 1), \\ h &= \operatorname{argmin}_k S_k \geq h', h' = \min(l' + 1, S_{n_e}), \\ \gamma &= \eta - \lfloor \eta \rfloor \end{aligned} \quad (13)$$

where argmin_k is a function denoting the minimum of the expression over k and $\lfloor \eta \rfloor$ is the ‘floor’ function returning the largest integer smaller than η .

As a simple illustrative example, suppose $q = 0.5$, $n_e = 10$ and each $w_i = 1$. This implies $S_1 = 1, S_2 = 2, \dots, S_{n_e} = n_e$ and $\eta = 5.5$. Hence, $l = 5$, $h = 6$, $\gamma = 0.5$ and the estimate of the 0.5 quantile is just the average of the 5th and 6th order statistics. This makes intuitive sense because these are just the ‘middle’ values in the sample. Loosely speaking η is close to $100q\%$ of S_{n_e} and the order statistics that are averaged are those associated with the S_k closest to (and larger than) the integers above and below η .

2.5.3 Estimating Sampling Uncertainty in Quantile Estimates

The bootstrap procedure described in Section 2.2 above can be used to construct sampling confidence intervals on the epistemic quantiles. In the context of xLPR, a 90% confidence interval on the 95th epistemic quantile has the interpretation: ‘there is 90% confidence that 95% of the population of epistemic samples falls below this bound.’

For example, to compute an upper confidence interval for the 95th quantile of the response, the bootstrap algorithm in Section 2.2.2.2 can be directly applied by using the 95th quantile of the bootstrapped epistemic samples as the QoI in step 2. For a $1 - \alpha$ confident upper bound on this quantity, the $1 - \alpha$ quantile of the estimated 95th quantile is computed elementwise. Such a bound has the interpretation: ‘there is $100(1 - \alpha)\%$ confidence that 95% of all epistemic responses fall below this bound.’

Estimating quantiles of low probabilities is typically challenging because enough variability must be observed in the probabilities over epistemic samples to calculate quantiles. Quantile estimation also becomes more challenging as quantiles become more ‘extreme,’ i.e., further out in the tails (0.95, 0.99, 0.999). Extreme quantiles are difficult to estimate and require much larger sample sizes to achieve accurate estimates. Example sampling confidence intervals for the 95th percentile of the probability of rupture are shown in Figure 9 below for the 1000/50 SRS with importance sampling. Accurately estimating percentiles requires more samples than accurately estimating means. The step-like nature of the quantile estimates is shown clearly in Figure 9. With enough samples, the estimates would be a smooth function of time. As an example, when estimating the mean (as in Figure 7 and Figure 8), the estimate is an essentially smooth function over time because smaller sample sizes are acceptable for estimation of the mean.

It is important to emphasize that, when trying to estimate epistemic quantiles, more samples and more targeted importance sampling will be required than when simply trying to estimate a mean (the ‘best-estimate’) over all aleatory and epistemic samples (as in the scenario analysis). Choosing an epistemic sample size that is too small will result in underestimation of epistemic uncertainty in the QoI. Increasing the aleatory sample size will also be required for accurate estimation of epistemic quantiles and their corresponding confidence intervals, as discussed in Section 4.1.

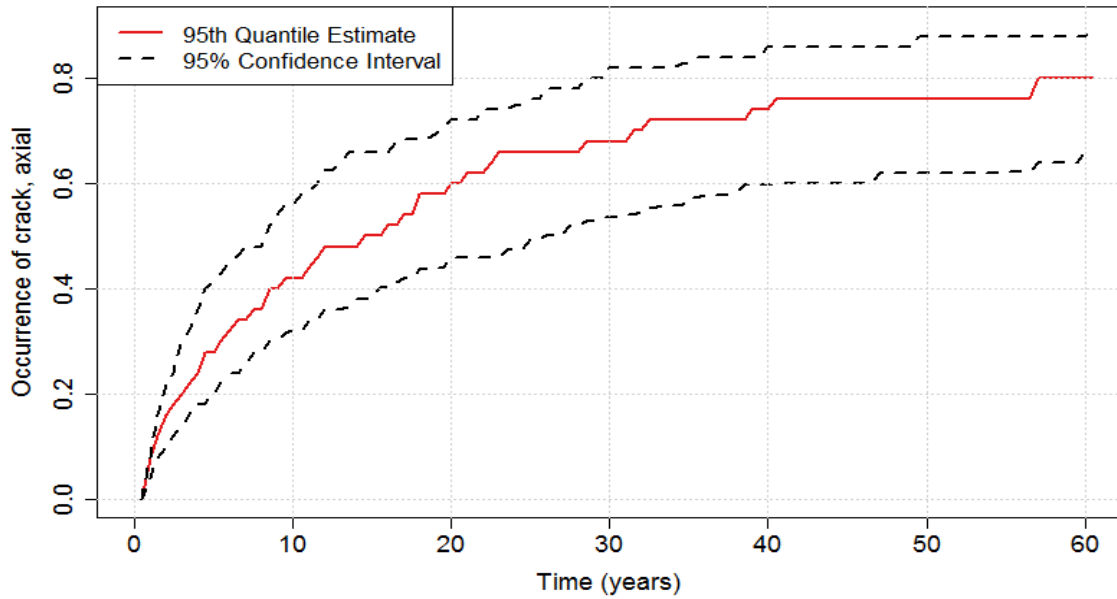


Figure 9: Sampling confidence intervals on the 95th percentile of the epistemic realizations for 1000/50 SRS with IS.

2.5.4 Visualizing Uncertainty Analysis Results

UA results can be visualized using plots of the QoIs over epistemic samples. For example, the estimated probabilities of occurrence of axial crack using results under Scenario 3 is displayed in Figure 10. This run (1000 epistemic/50 aleatory) was used as it has sufficiently small variability as determined by the Convergence Analysis in Section 2.4.

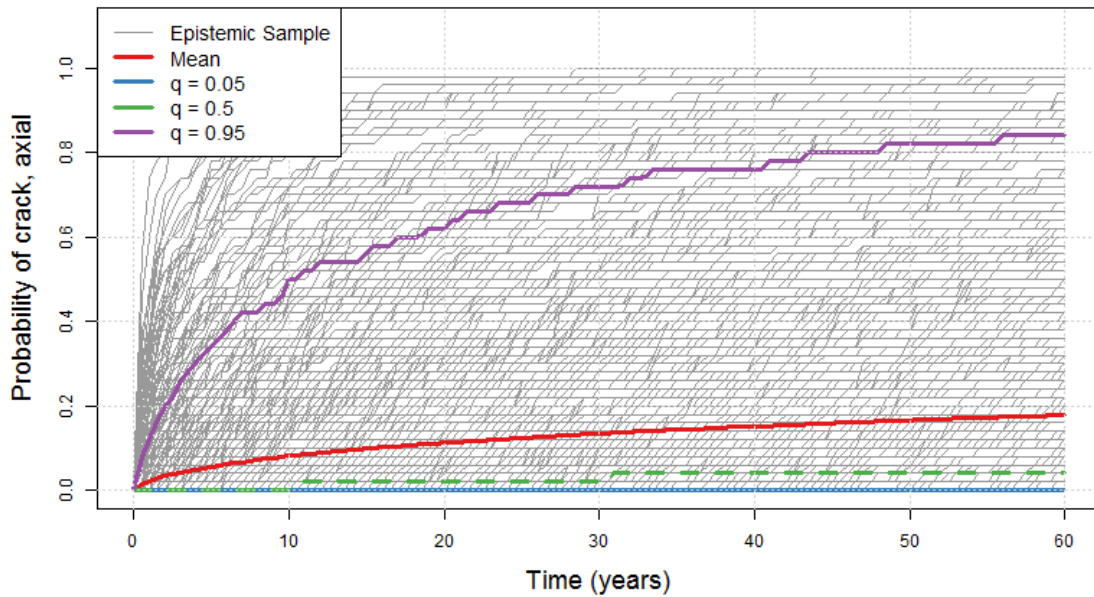


Figure 10: Probability of occurrence of axial crack for each epistemic realization (grey), the mean (red), and the 0.05 (blue), 0.5 (green), and 0.95 (purple) quantiles.

The figure shows the probability of occurrence of axial crack for each epistemic realization (grey lines), the mean (red line), and the 0.05, 0.5, and 0.95 quantiles (blue, green, and purple lines, respectively). The quantiles shown here offer a summary of the total range of results. Plots such as this can be used to gain a better understanding of the spread of epistemic realizations and to analyze quantiles of interest. For example, the 0.5 quantile for axial crack is about 0.05 at year 60, meaning that half of the epistemic realizations have larger than a 5% chance of a crack by year 60. One can also gain a better understanding of the relationships between different outputs by examining such quantiles. For example, Figure 11 displays the mean (red) and 0.95 quantiles (purple) for axial crack (dashed) and leak (solid). The intuitive result that cracks occur before leaks is confirmed in the plot as the mean/0.95 quantiles of leak trails behind that of crack.

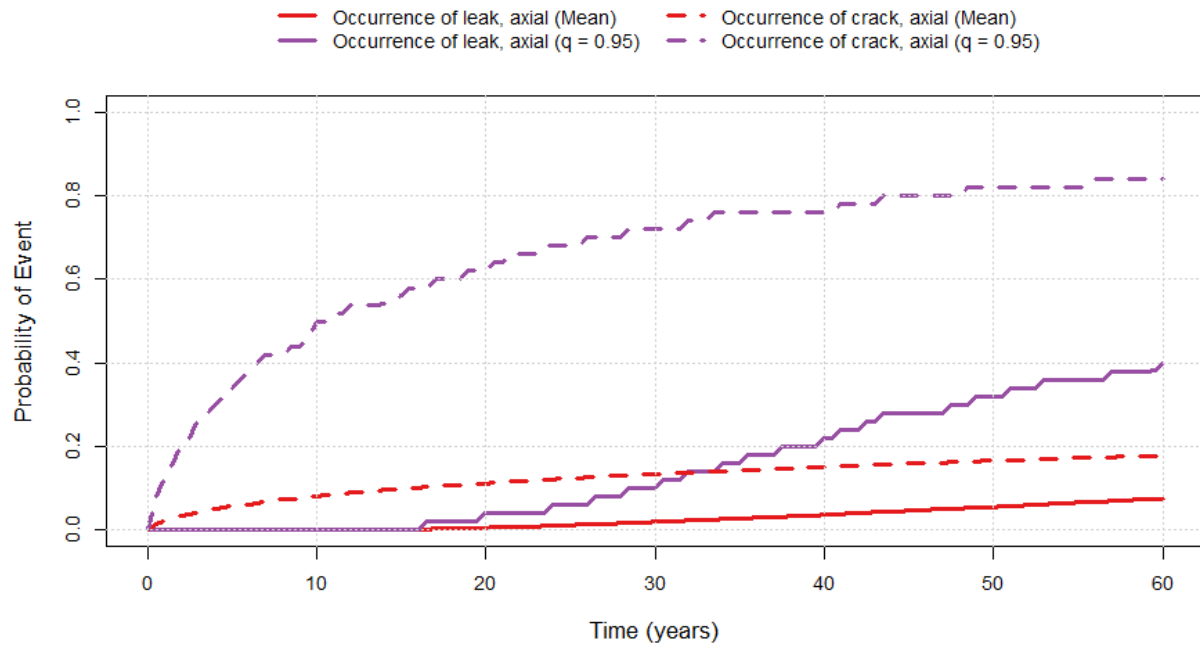


Figure 11: Mean probability (red) and 0.95 quantile (purple) for occurrence of axial leak (solid) and crack (dashed).

This page intentionally left blank.

3 SCENARIO-SPECIFIC RESULTS

The inputs for a set of 11 scenarios were developed by the xLPR Inputs Group. The defining features for each of the scenarios are given in Table 2. All of these scenarios examine the behavior of a reactor pressure vessel (RPV) outlet weld during a 60-year simulation. The analysis of these scenarios in this report is limited to a subset of the outputs available in the xLPR Version 2.0 code. However, the methods presented in this report could be used to analyze other xLPR outputs as needed in future applications. The inputs used for each of these 11 scenarios will be available on the xLPR portal for future reference.

Table 2: Defining features for the 11 xLPR scenarios.

Scenario	Initiation	Growth	Flaw Orientation	Mitigation	
				Type	Timing
Scenario 1	Initial Flaw	Fatigue	Circumferential	None	---
Scenario 2	PWSCC	PWSCC	Circumferential	None	---
Scenario 3	PWSCC	PWSCC	Circumferential & Axial	None	---
Scenario 4	PWSCC	PWSCC	Circumferential & Axial	MSIP	20 Years
Scenario 5	PWSCC	PWSCC	Circumferential & Axial	MSIP	40 Years
Scenario 6	PWSCC	PWSCC	Circumferential & Axial	Zn	20 Years
Scenario 7	PWSCC	PWSCC	Circumferential & Axial	H2	20 Years
Scenario 8	PWSCC	PWSCC	Circumferential & Axial	Zn & H2	20 Years
Scenario 9	PWSCC	PWSCC	Circumferential & Axial	Inlay	40 Years
Scenario 10	PWSCC & Fatigue	PWSCC & Fatigue	Circumferential & Axial	MSIP, Zn, & H2	20 Years
Scenario 11	Fatigue	Fatigue	Circumferential & Axial	None	---

Scenarios 1 and 11 utilize only fatigue growth and experienced very different results than the remaining scenarios. Scenarios 2 through 10 are defined by primary water stress corrosion cracking (PWSCC) initiation and growth options. Scenario 3 is used as a comparative base case for Scenarios 2 and 4-10. This scenario also served as an illustration of the simulation workflow developed for the purposes of the scenario analysis effort. Because of this, the description of the results of Scenario 3 contains many additional studies and analyses that are used to draw larger conclusions about sampling selection and are referenced throughout this report.

The sections that follow describe the results and conclusions of the analysis of each of the 11 xLPR scenarios. These results are dependent on the particular inputs selected for each scenario. These inputs were selected by the xLPR Inputs Group. Input parameters discussed throughout this chapter are identified based on their alphanumeric identification code (i.e., p####) from the xLPR input set and by their parameter description. Runs utilizing importance sampling (IS) are distinguished by their selected target quantile level (i.e., 0.95). This target quantile level determines the location about which importance sampling will take place.

The simulation results presented in this chapter were run over an extended period of time that spanned several beta releases of the xLPR code. Because of this, the results presented in this chapter

were run using both xLPR-2.0 Beta2.1_R12 and xLPR-2.0 Beta2.1_R13 with GoldSim version 11.1.5. The changes made to the xLPR code between these beta releases did not impact the results of the calculations for these scenarios. Test cases were run to confirm that the results would be the same across versions.

Several additions, as described below, were made to the beta versions of the xLPR V2.0 code in order to generate the data needed to analyze each scenario. These additions were made using the GoldSim version control features and do not impact the calculations or results of the xLPR V2.0 code. The changes made inside of the ‘Main_Model’ element are shown in Figure 12. The model tree shown on the left of this figure displays changes made when compared to the original code in red.

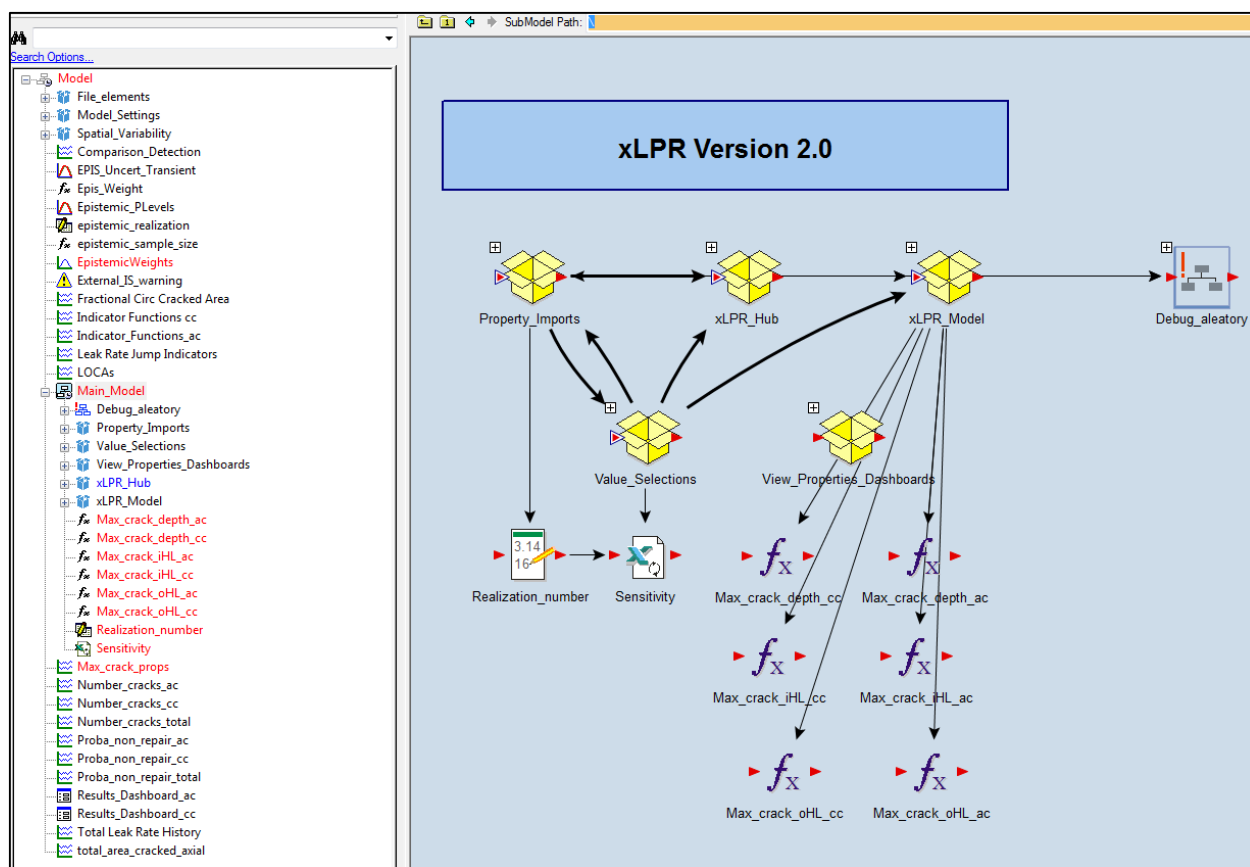


Figure 12: Additions made to the xLPR model required for the Scenario Analysis report calculations (right). Version control changes are shown in red (left).

First, the sampled values for all inputs used in the code Scenario Analysis input sets were added as outputs to a GoldSim ‘Spreadsheet’ element located inside of the ‘Main_Model’ container. The overall realization number was calculated as a function of the epistemic and aleatory realization numbers. This allowed the sampled values to be exported to a spreadsheet for each aleatory realization using the overall realization number as the export row. This implementation was tested by setting each of these values to constant with a magnitude of their input number in the xLPR input set. Inspection showed that all sampled values were being exported to the expected columns in the final spreadsheet.

Second, additional outputs were added related to the calculation of crack properties. The maximum crack depth, inner half-length, and outer half-length for both circumferential and axial cracks over all subunits at each timestep was calculated at the aleatory level inside of the 'Main_Model' container. These results were added to the 'Main_Model' interface so that they could be extracted at the epistemic level and result history elements for these outputs were added at the model root so that they would be saved for all epistemic realizations.

Finally, several additional changes were made at the model root. A GoldSim distribution result element was added at the model root to save the value of the epistemic sampling weight for each epistemic realization. This weight is required in order to calculate result statistics, especially when importance sampling is applied. Crack-specific results for cracks 1 through 5, included in the xLPR code by default, were removed from the 'Main_Model' element interface and their related GoldSim result history elements were deleted in order to save memory so that scenario simulations could run to completion.

All of the results calculated for each scenario were extracted from their respective GoldSim result elements and saved in an Excel spreadsheet for each simulation. This spreadsheet is read by the post-processing R code to generate the results and plots that are provided for each scenario in the following sections.

3.1 Scenario 1

3.1.1 Scenario Summary

Scenario 1 examines fatigue flaw growth on an RPV outlet weld. Initial flaws are assumed and PWSCC growth is turned off so that cracks only grow by fatigue.

Scenario 1 uses the following defining options in the xLPR code:

- Initial flaw
- Circumferential flaw orientation
- Fatigue flaw growth
- No mitigation

The outputs of interest defined for this scenario for circumferential cracks are:

- Maximum crack depth
- Maximum crack inner half-length
- Maximum crack outer half-length

Under the transients defined for this scenario, crack leaks and ruptures do not occur. Because initial flaws are assumed, the probability of crack occurrence is always equal to 1 after the first timestep.

The simulations performed for this scenario are listed below.

1. 100/50 (epistemic/aleatory sample size) with SRS
2. 500/50 with SRS
3. 1000/50 with LHS
4. 1000/50 with SRS
5. 1000/50 with LHS and importance sampling on input parameter p9001 (Fatigue Growth Threshold SIF Scaling Factor, CKTH) with target quantile of 0.05
6. 1000/50 with LHS and importance sampling on input parameter p9001 (Fatigue Growth Threshold SIF Scaling Factor, CKTH) with target quantile of 0.95

Scenario 1 results in very little crack growth and low variability across all realizations, even for simulations using importance sampling. Convergence testing indicates that this result is stable and reproducible regardless of random seeds and sample sizes. Using the current Scenario 1 input set, any of the sampling options shown above could be used to produce the same mean crack properties.

The following section describes the progression of simulations used to analyze Scenario 1. Subsequent sections describe the overall comparison between these simulations, epistemic uncertainty analysis of the Scenario 1 results, and a convergence analysis on one of the sampling schemes used for this scenario.

3.1.2 Analysis Progression

The run-to-run comparison for all of the sampling schemes examined in the analysis of Scenario 1 is described in this section. The maximum circumferential crack inner half-length is used to exemplify the results of these comparisons because this result was most strongly impacted by the differences in sampling schemes that differentiate each run. However, the same conclusions presented below can be drawn using comparisons of the maximum circumferential crack depth for this scenario.

Runs 1 and 2 assess the impact of changes to epistemic sample size on result uncertainty. Figure 13 plots the mean maximum circumferential crack inner half-length for Run 1 and Run 2 along with the 95% confidence intervals (CIs). These CIs describe the uncertainty associated with the mean due to a finite sample size. Figure 14 plots the width of the CIs for both runs 1 and 2 through time. The maximum circumferential crack inner half-length shown is unitless and represents the inner half-length normalized by the inner circumference of the pipe for each realization. This normalization is important to note when examining the results shown below; the mean circumferential inner half-length does not grow larger than approximately 3.75% of the inner pipe circumference, a value that is not much greater than the mean initial flaw inner half-length. Because these numbers are very small and their variation from realization to realization is also small, the results shown must be considered in the context of their relative scale.

Although the means are nearly the same for these two simulations, the confidence interval comparison indicates that an increase from 100 to 500 epistemic samples leads to a reduction in the uncertainty associated with the mean maximum circumferential crack inner half-length although, as noted above, this reduction is very small.

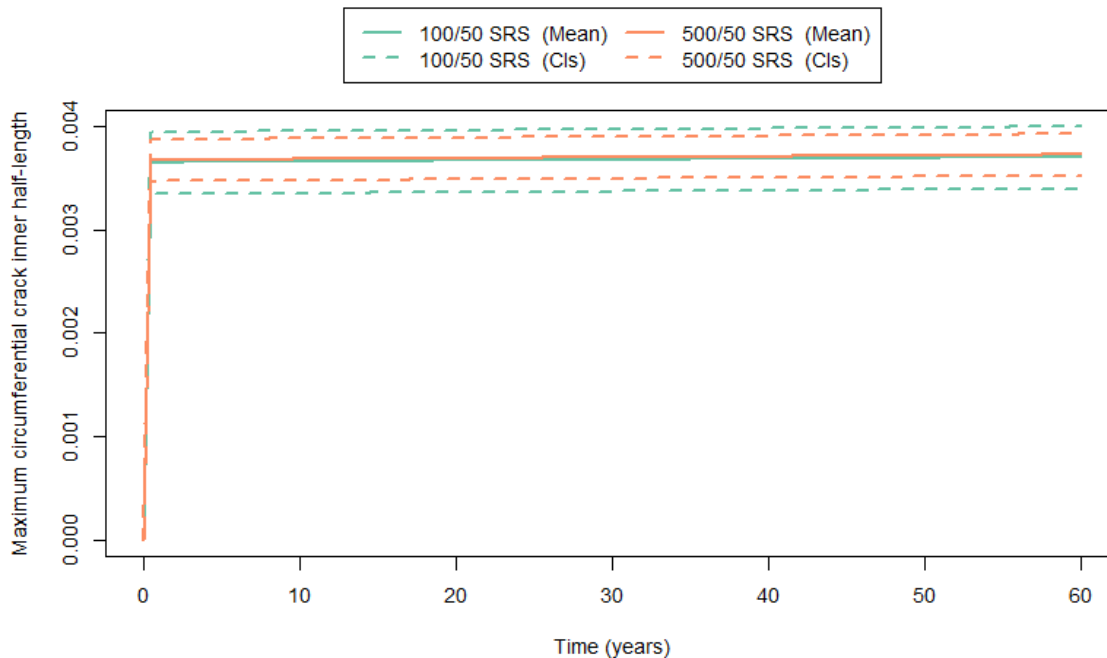


Figure 13: Mean maximum circumferential crack inner half-length (solid line) and 95% confidence intervals (dashed lines) for Scenario 1, Run 1 (green) and Run 2 (orange).

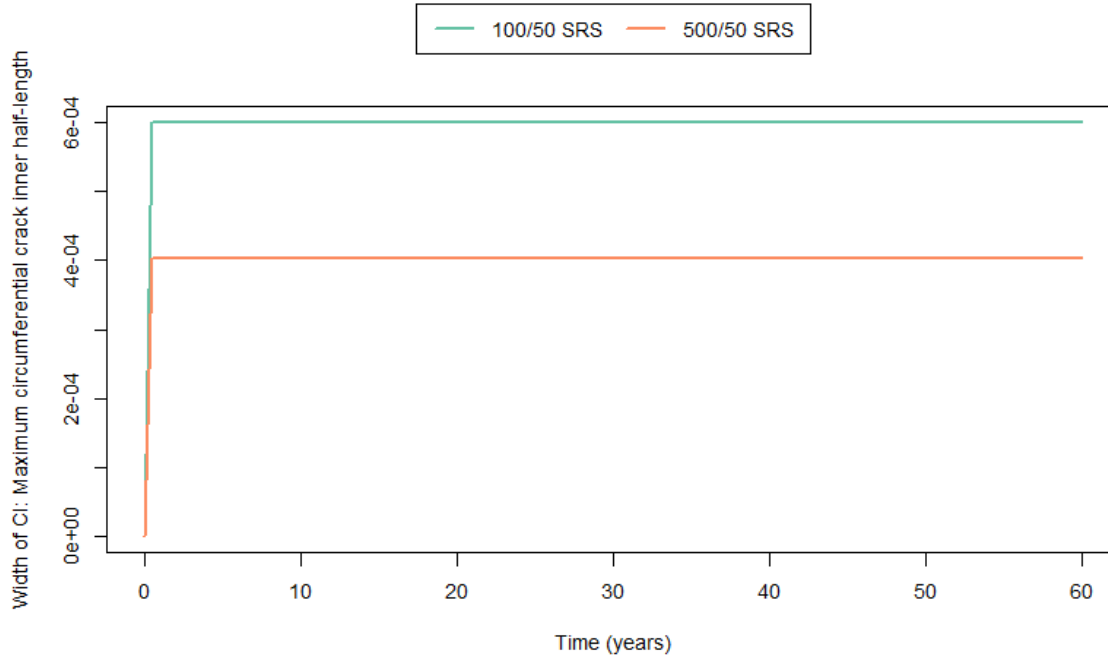


Figure 14: Width of 95% confidence interval for the mean maximum circumferential inner half-length for Scenario 1, Run 1 (green) and Run 2 (orange).

Runs 3 and 4 use 1000 epistemic samples and LHS and SRS, respectively. The reduction in uncertainty in the mean maximum circumferential crack inner half-length is even smaller when moving from 500 to 1000 epistemic samples than it was when transitioning from 100 to 500 samples, as is shown in Figure 15. LHS is expected to decrease the width of confidence intervals about the mean more than SRS, as is described in Section 4.3. This method was applied along with an epistemic sample size of 1000 for Run 3.

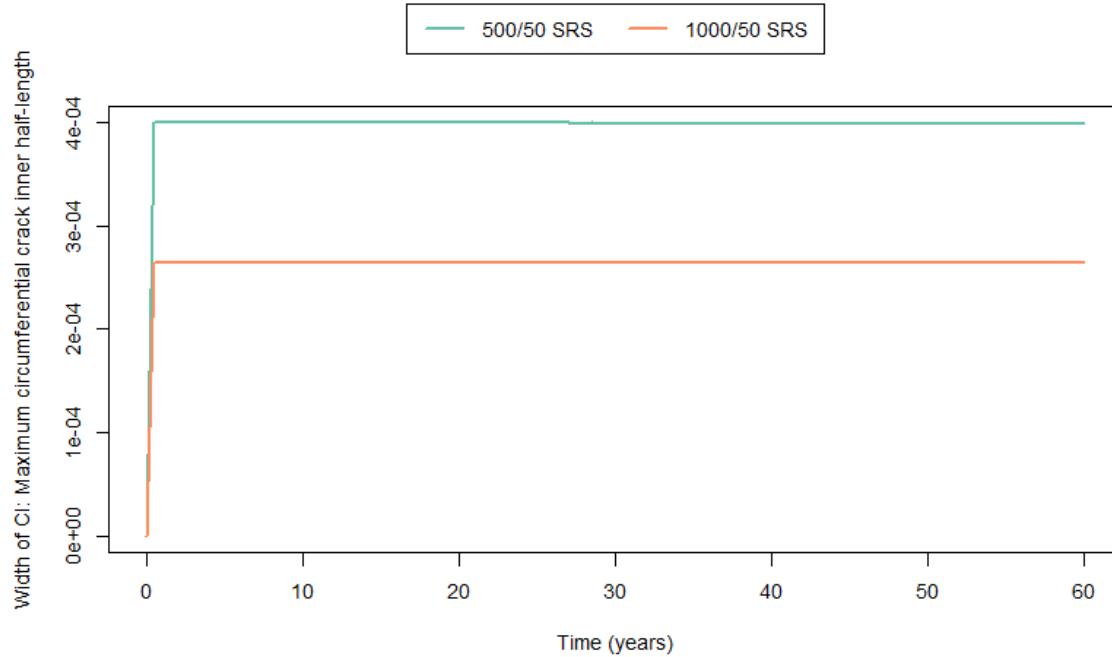


Figure 15: Width of 95% confidence interval for the mean maximum circumferential crack inner half-length for Scenario 1, Run 2 (green) and Run 4 (orange).

A sensitivity analysis was conducted for Run 3 to determine if any of the uncertain input parameters for this scenario could be identified for importance sampling. A stepwise rank regression analysis was used to determine how the uncertainty in the input variables relates to the uncertainty in the outputs for Scenario 1. The results of this rank regression analysis, shown in Table 3, compare the standardized rank regression coefficients (SRRCs) for the five most significant variables found in Run 3 for each output.

This sensitivity analysis was convoluted by the necessary dependence between the pipe thickness and the outputs of interest; the maximum crack depth and inner half-length calculated in this analysis are normalized by the pipe thickness and inner pipe circumference (a function of pipe thickness). In order to understand how the inputs and outputs are related when this dependence is removed, the absolute crack depth was added as an output of interest for this scenario. This output was calculated as a post-processing step outside of the simulation using the sampled pipe thickness and normalized crack depth. The results of the sensitivity analysis of this output are also shown in Table 3.

Table 3: Summary of the stepwise rank regression analysis results for all outputs at year 60 for Scenario 1, Run 3.

Variable Identifier	Variable Name	SRRC
Maximum Circumferential Crack Depth – $R^2 = 0.941$		
p1102	Pipe Wall Thickness	-0.970
p6001	Transient #1 Uncertainty Multiplier	0.074
p9001	Fatigue Growth CKTH	-0.039
p4352	Axial WRS Pre-mitigation	-0.033
p6003	Transient # 3 Uncertainty Multiplier	0.026
Maximum Circumferential Inner Half-Length – $R^2 = 0.013$		
p5108	b (axial)	0.072
p2508	Material Init J-Resist Exponent, m	-0.071
p6005	Transient # 5 Uncertainty Multiplier	0.070
p1102	Pipe Wall Thickness	0.061
p5103	Intercept, B0 (axial)	-0.060
Absolute Circumferential Crack Depth – $R^2 = 0.155$		
p6001	Transient # 1 Uncertainty Multiplier	0.277
p1102	Pipe Wall Thickness	-0.167
p9001	Fatigue Growth CKTH	-0.121
p6003	Transient # 3 Uncertainty Multiplier	0.107
p4352	Axial WRS Pre-mitigation	-0.102

The results above show that the sampled pipe thickness is almost perfectly negatively correlated with the maximum circumferential crack depth. This result is intuitive because the maximum crack depth is normalized by, and is therefore a function of, the pipe thickness; the thicker the pipe becomes, the smaller the normalized crack depth becomes. This result is exacerbated by the small crack growth shown in this scenario. Subsequent scenarios with much more significant growth do not experience the same dependence between maximum crack depth and thickness as other factors contributing to faster growth become more important.

The R^2 value for the maximum circumferential inner half-length stepwise rank regression are very low and the top five SRRC values for this regression are also extremely small. Because of this, these regression results were not used to inform importance sampling decisions for this scenario. The R^2 value for the maximum absolute circumferential crack depth is also very small but, in order to explore sampling options to the greatest extent possible for Scenario 1, these results were used to inform an importance sampling selection. The ability to apply importance sampling to p6001(Transient #1 Uncertainty Multiplier) is not available in the current version of xLPR and importance sampling p1102 (pipe wall thickness) would not provide insight into the fatigue growth mechanisms under analysis in this scenario. These considerations led to the selection of p9001(Fatigue Growth Threshold SIF Scaling Factor, CKTH) for importance sampling. Examination of the scatterplot of this variable related to the output of interest motivated the selection of a quantile of 0.05 for importance sampling as smaller values of this variable led to a few larger absolute crack depths. Figure 16 shows a scatterplot of this variable and the absolute crack depth for Run 3. This scatterplot shows that the dependence of absolute crack depth on p9001 is very weak. The same relationship is shown for Run 5 in which importance sampling is applied. This scatterplot shows that while some larger absolute crack depths are generated, the majority of results calculated under importance sampling remain close to the mean result behavior.

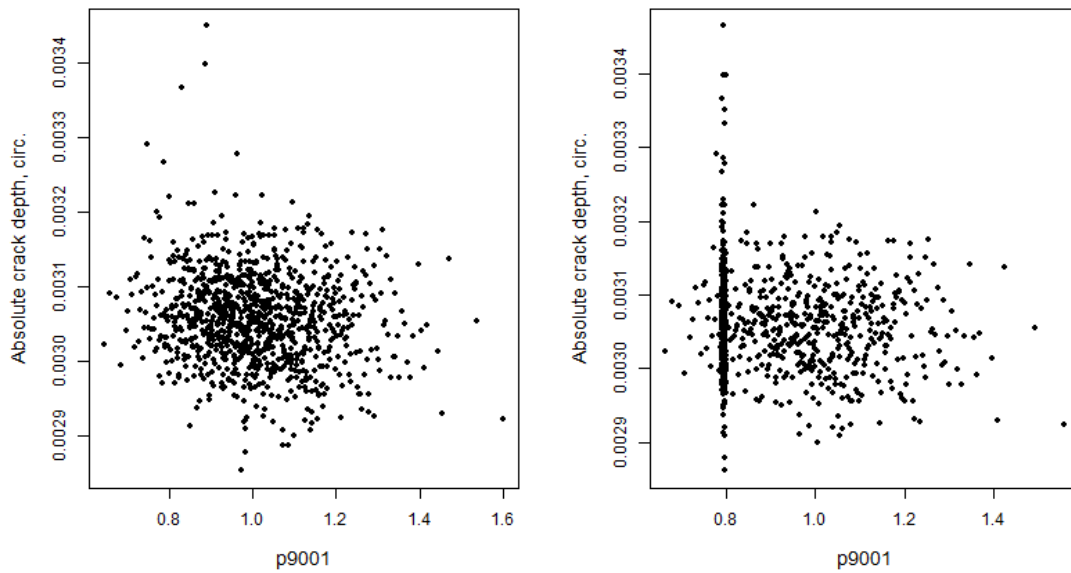


Figure 16: Scatter plots of p9001 (Fatigue Growth Threshold SIF Scaling Factor, CKTH) and the absolute crack depth for Scenario 1, Run 3 (left) and Run 5 (right).

The width of the confidence intervals about the mean results do not shift by a significant amount using importance sampling, as shown in Figure 17.

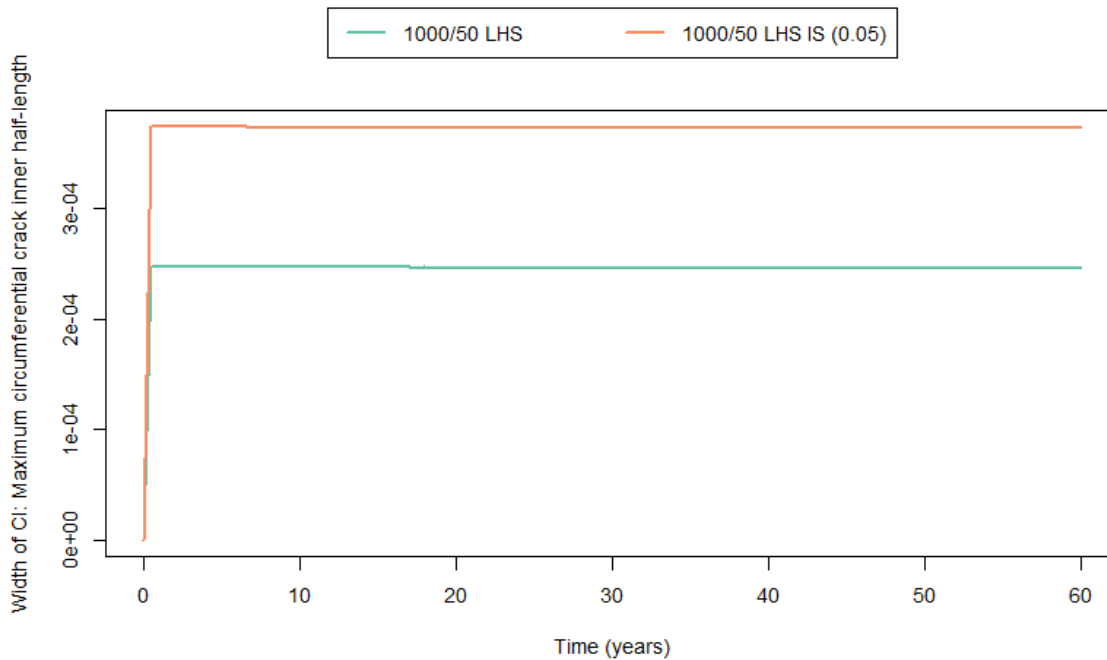


Figure 17: Width of 95% confidence interval for maximum circumferential crack inner half-length for Scenario 1, Run 3 (green) and Run 5 (orange).

Low values of variable p9001 (Fatigue Growth Threshold SIF Scaling Factor, CKTH) are expected to produce smaller cracks because this variable acts as a scaling factor for the threshold that is used for determining whether transient stresses will be used in fatigue growth calculations. The negative SRRC for this variable also predicts this behavior. However, for the sake of completeness and to determine if this expectation was correct, importance sampling on variable p9001 was also run using a higher quantile of 0.95. This importance sampling did not yield additional information about the outputs of interest because results within the importance sampling interval span the range of the original values as shown in Figure 18.

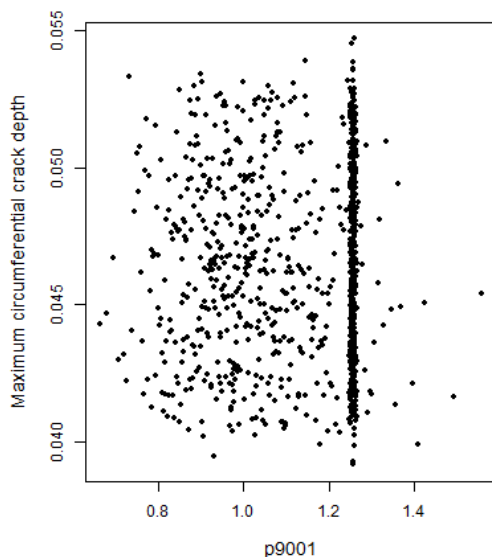


Figure 18: Scatterplot for p9001 (Fatigue Growth Threshold SIF Scaling Factor, CKTH) and maximum circumferential crack depth for Scenario 1, Run 6.

This importance sampling selection also did not impact the width of confidence intervals or the mean result for the outputs of interest.

Importance sampling did not provide any additional information to the characterization of the uncertainty in the outputs of interest for this scenario. This may be a function of the variable selected for importance sampling; the variable that might have been better for this selection (p6001, Transient #1 Uncertainty Multiplier) is not able to be importance sampled in the current version of the xLPR code. Any additional importance sampling selections are not supported by the results of the sensitivity analysis shown in Table 3.

3.1.3 Overall Results

The analysis of Scenario 1 was somewhat limited because of the small amount of crack growth caused by fatigue under the transients and stresses defined for this scenario. The results showed very little epistemic uncertainty and experienced little sampling variability between scenario runs. Future exploration of this scenario could include increased transient loads and transient frequencies. These fatigue input changes are further explored in the analysis of Scenario 11 described in Section 3.11.

Figure 19 shows that the mean maximum circumferential crack depth remained the same across all of the sampling options examined for Scenario 1. The mean maximum circumferential inner half-length was also nearly the same across all sampling options (Figure 20) although, as described in Section 3.1.2, the 95% CIs about the mean do decrease when 1000 epistemic samples under LHS are used. The importance sampling results did not yield any significant reduction in the uncertainty associated with the mean results. Much of the variation across all results is due to the sampled initial flaw geometry; calculations subsequent to this did not strongly influence result variability.

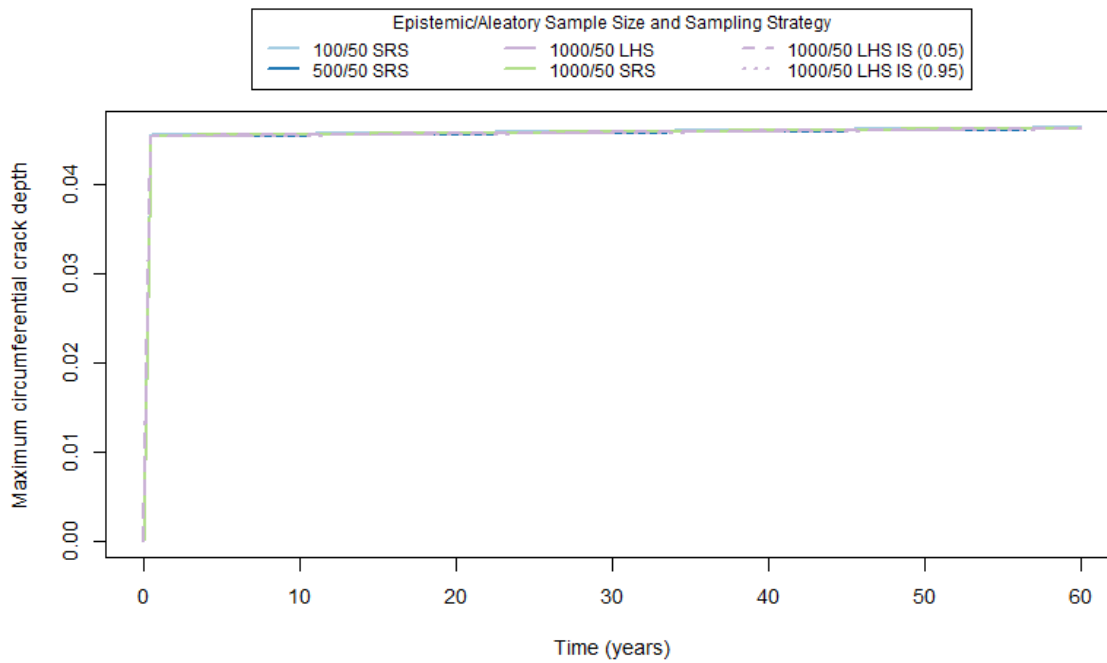


Figure 19: Mean maximum circumferential crack depth for Scenario 1, Runs 1 through 6.

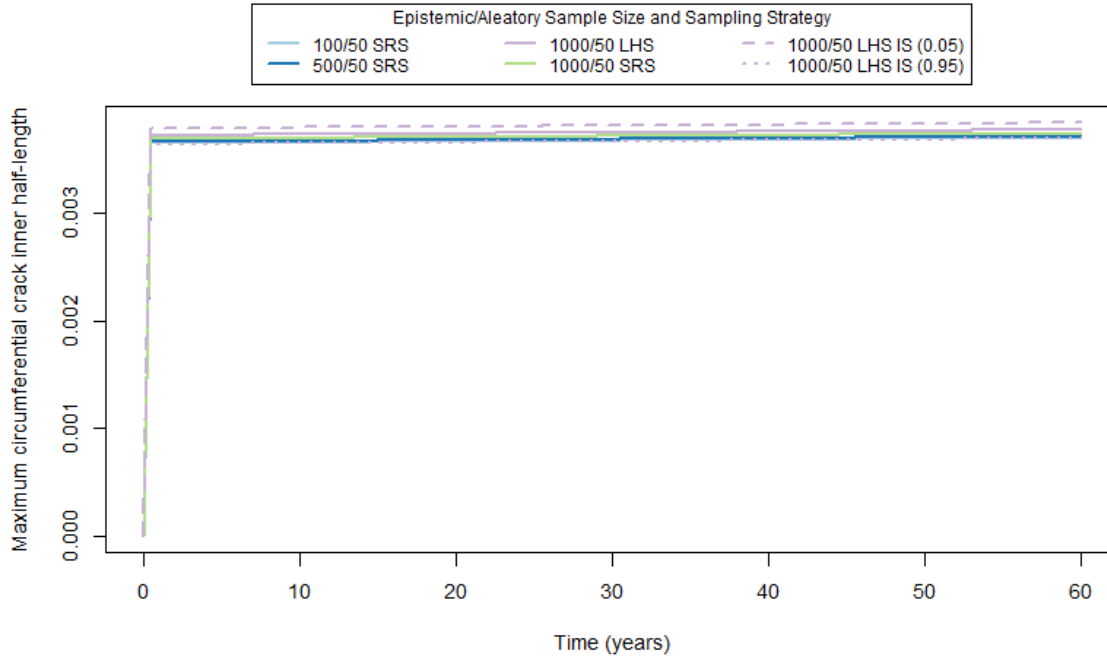


Figure 20: Mean maximum circumferential crack inner half-length for Scenario 1, Runs 1 through 6.

3.1.4 Epistemic Uncertainty Analysis

The epistemic uncertainty analysis was conducted using a sampling scheme for Scenario 1 of 1000 epistemic and 50 aleatory samples and LHS without IS (Run 3). Figure 21 shows a plot of the maximum crack inner half-length for each epistemic realization (grey), the mean (red), and the 5th (blue), 50th (green), and 95th (purple) percentiles over time. Figure 22 shows the same result for maximum crack depth.

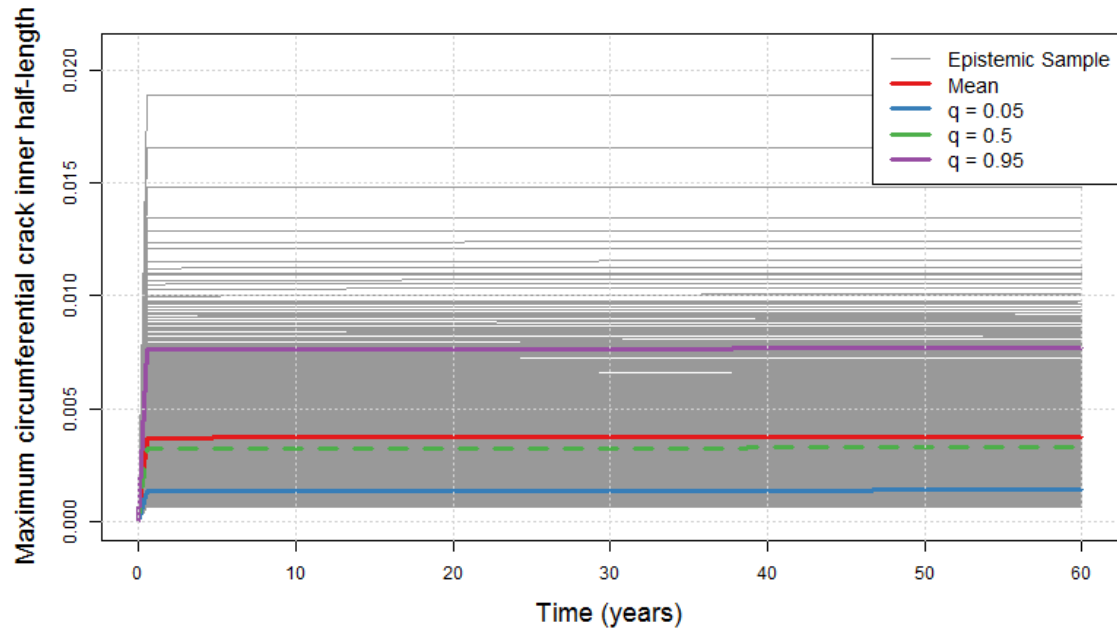


Figure 21: Maximum circumferential crack inner half-length for each epistemic realization (grey), the mean (red), and the 5th (blue), 50th (green), and 95th (purple) percentiles for Scenario 1, Run 3.

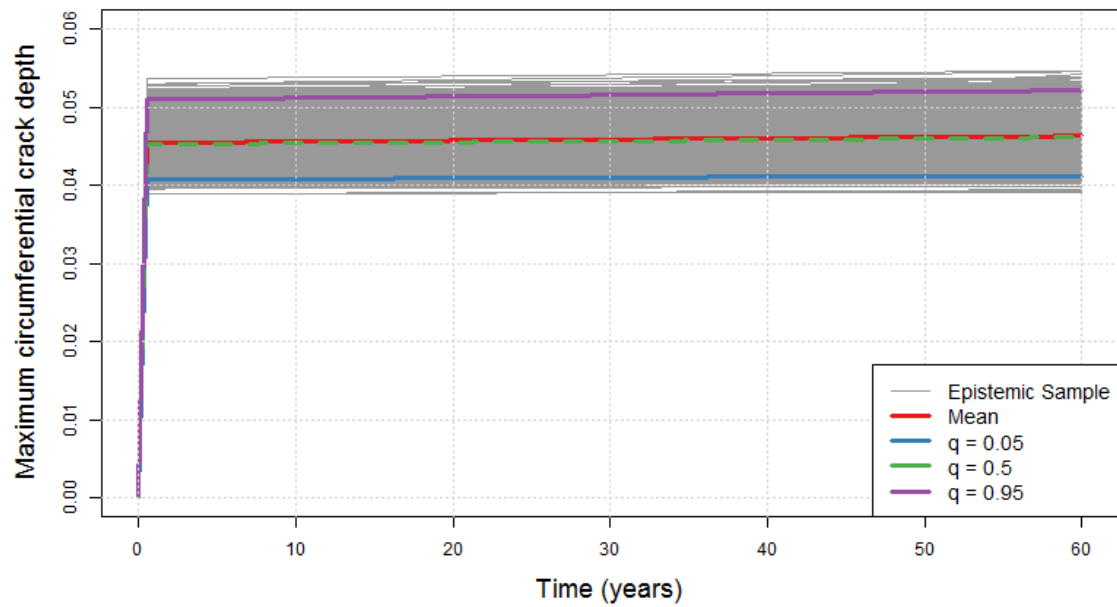


Figure 22: Maximum circumferential crack depth for each epistemic realization (grey), the mean (red), and the 5th (blue), 50th (green), and 95th (purple) percentiles for Scenario 1 Run 3.

The epistemic uncertainty analysis results show that 95% of the epistemic realizations have a maximum crack inner half-length less than 0.0075 and a maximum crack depth less than 0.05. In the context of this scenario, “extreme” cracks are still very small and do not threaten pipe stability.

3.1.5 Convergence Analysis

A convergence analysis was run to gain insight into the behavior of the sampling scheme for Run 3 (1000 epistemic and 50 aleatory samples and LHS without IS) using different random number seeds. This sampling scheme was re-run five times with different random number seeds for both the epistemic and aleatory loops. Each replicate was used to give an estimate of the mean maximum crack inner half-length. Figure 23 shows a comparison of these estimates. The five replicates were also used to estimate an overall mean value and a 95% prediction interval was created around this mean (Figure 24). The prediction intervals were constructed to determine the region that the next convergence run would be expected to fall within 95% of the time. These convergence results show that this sampling scheme is extremely stable as the means and prediction intervals are nearly identical. The same conclusions can be drawn using the maximum circumferential crack depth.

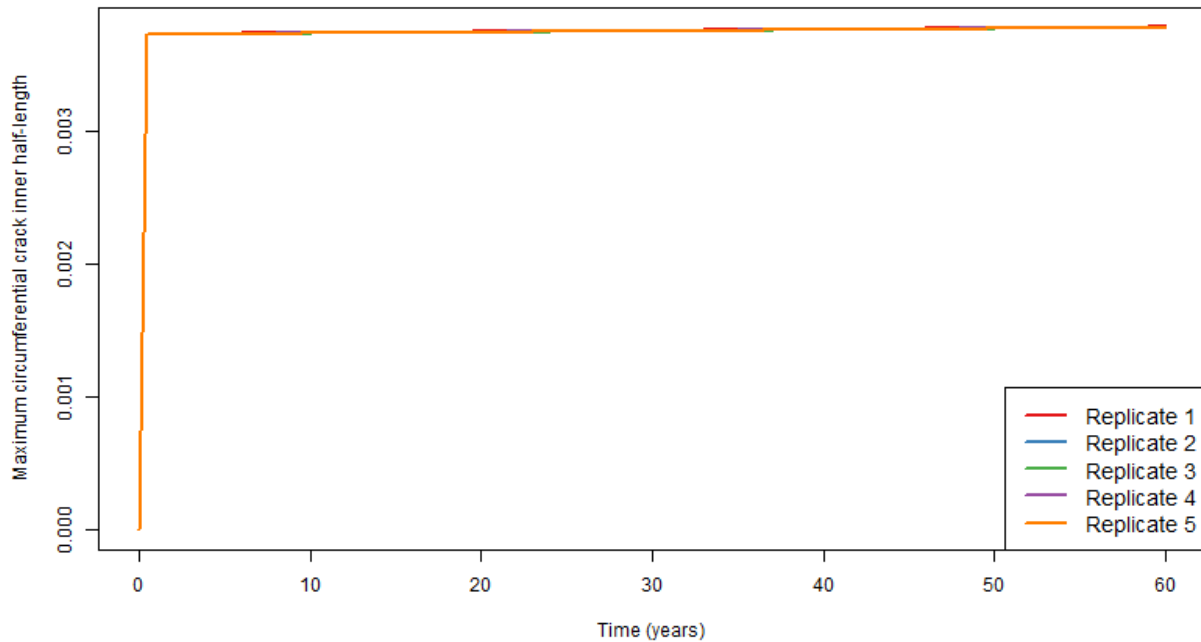


Figure 23: Comparison of estimated mean maximum circumferential crack inner half-length for five convergence replicates for Scenario 1, Run 3.

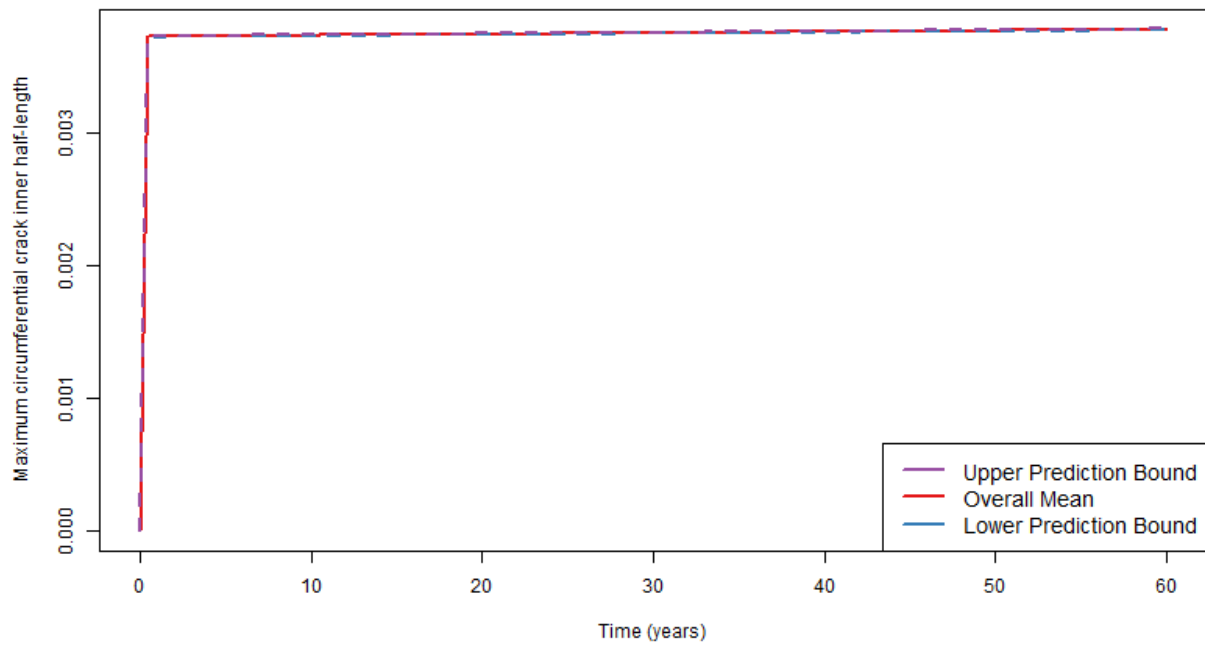


Figure 24: 95% prediction interval over mean maximum circumferential crack inner half-length for five replicates for Scenario 1, Run 3.

3.2 Scenario 2

3.2.1 Scenario Summary and Comparison to Scenario 3

Scenario 2 examines PWSCC flaw initiation and growth on an RPV outlet weld and uses the following defining options in the xLPR code:

- PWSCC flaw initiation
- Circumferential flaw orientation
- PWSCC flaw growth
- No mitigation or fatigue

This scenario is essentially the same as Scenario 3 (Section 3.3) except that it does not examine axial flaws.

The outputs of interest defined for this scenario for circumferential cracks are:

- Probability of occurrence of crack
- Probability of occurrence of leak
- Probability of occurrence of rupture
- Total Leak Rate
- Maximum crack depth
- Maximum crack inner half-length
- Maximum crack outer half-length

The comprehensive list of simulations that were used in the analysis of this scenario is shown below. A constant aleatory sample size of 50 was used for each of the sampling schemes considered.

1. 100/50 (epistemic/aleatory sample size) with simple random sampling (SRS)
2. 500/50 with SRS
3. 1000/50 with Latin Hypercube Sampling (LHS)
4. 1000/50 with LHS and importance sampling (IS) on p2543 (multiplier on direct model 1 (DM1) proportionality constant A) with target quantile of 0.95
5. 1000/50 with SRS and IS on p2543 with target quantile of 0.95

The results of Scenario 2 for circumferential cracks for the outputs of interest defined above are exactly the same as those found in Scenario 3, described in the following section (Section 3.3). This comparison is shown in Figure 25 through Figure 27 below. Because these results are identical and for the sake of brevity, the detailed examination of result uncertainty using the options selected for Scenario 2 will be left to the following description of the analysis of Scenario 3, described in Section 3.3.

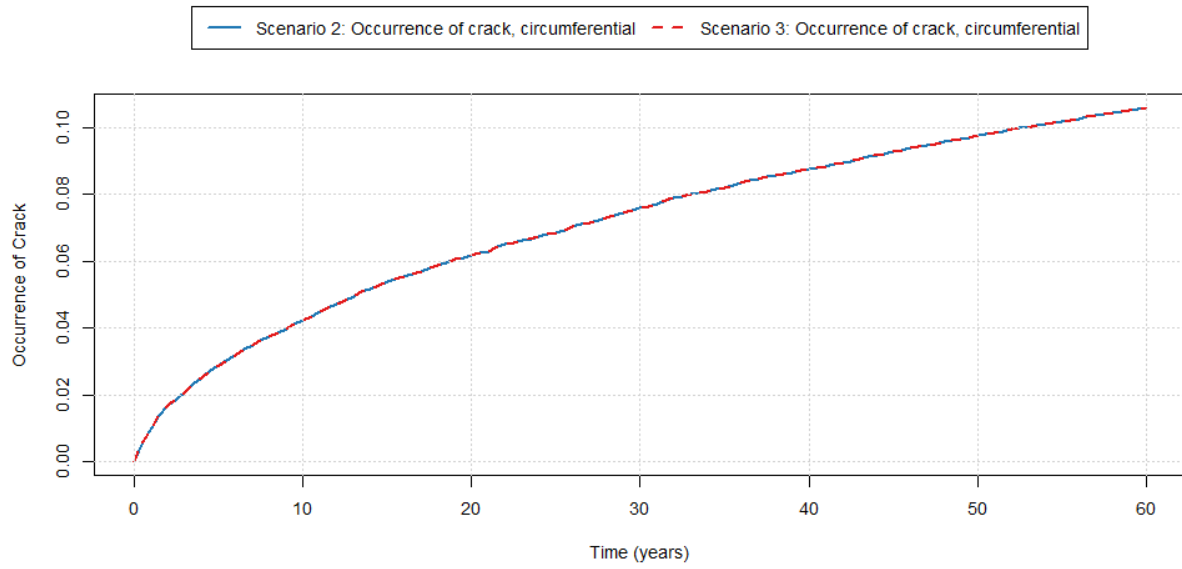


Figure 25: Mean probability of circumferential crack occurrence for Scenario 2 (blue) and Scenario 3 (red).

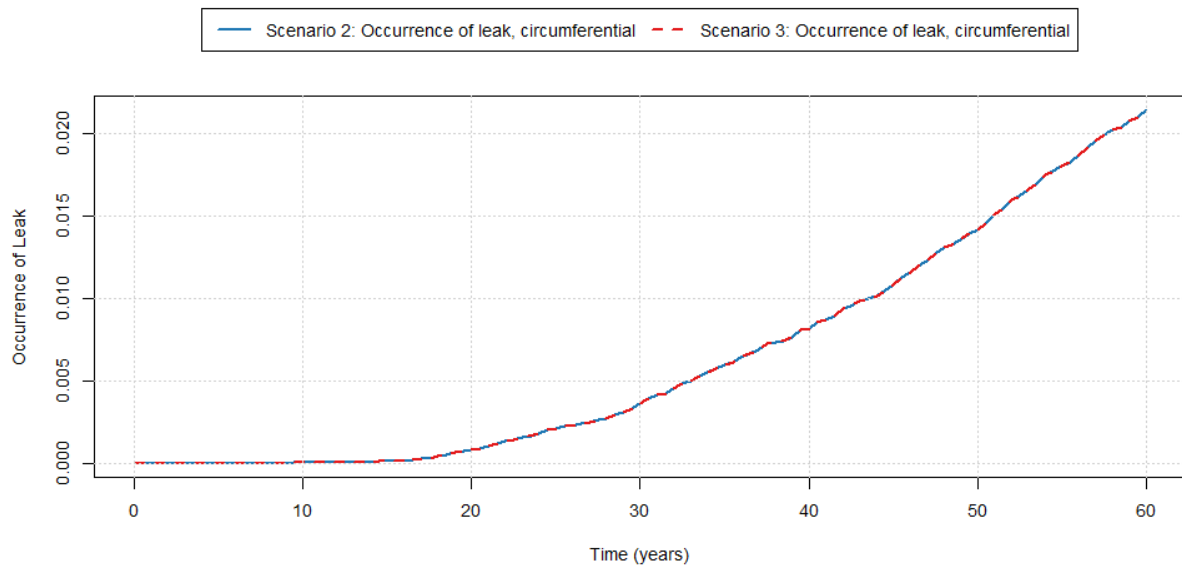


Figure 26: Mean probability of circumferential leak occurrence for Scenario 2 (blue) and Scenario 3 (red).

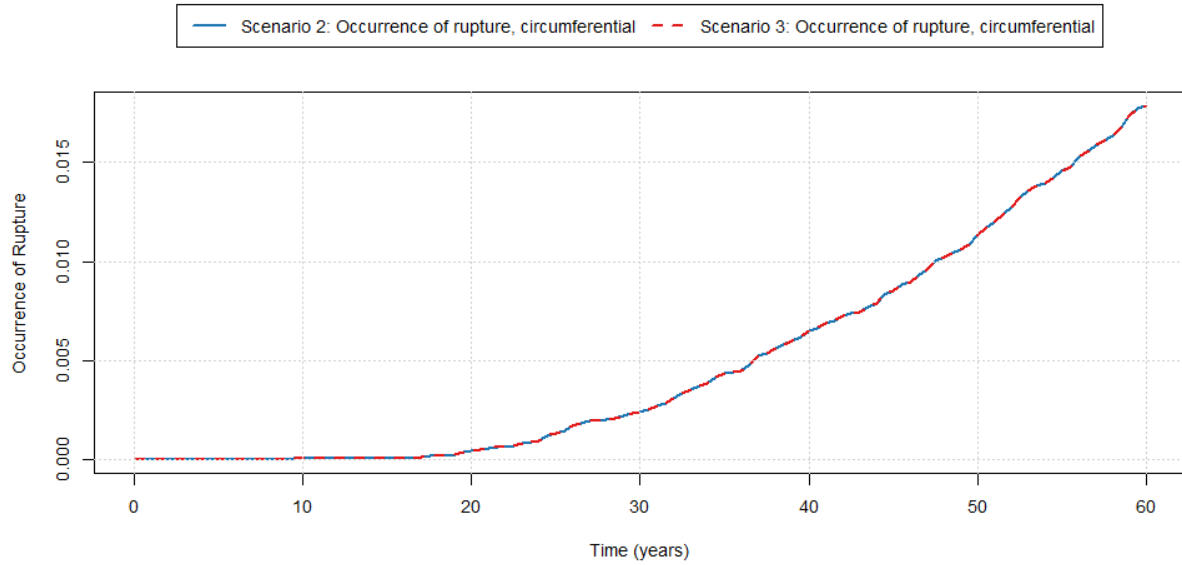


Figure 27: Mean probability of circumferential rupture occurrence for Scenario 2 (blue) and Scenario 3 (red).

Two outputs of interest, total leak rate and occurrence of rupture with ISI, are not the same for Scenario 2 when compared to Scenario 3. Total leak rate includes the leak rate from both circumferential and axial cracks and is thus expected to be lower when axial cracks are not included (note that axial leaks were observed for Scenario 3). The total leak rates [m^3/s] are compared for Scenario 2 and Scenario 3 in Figure 28.

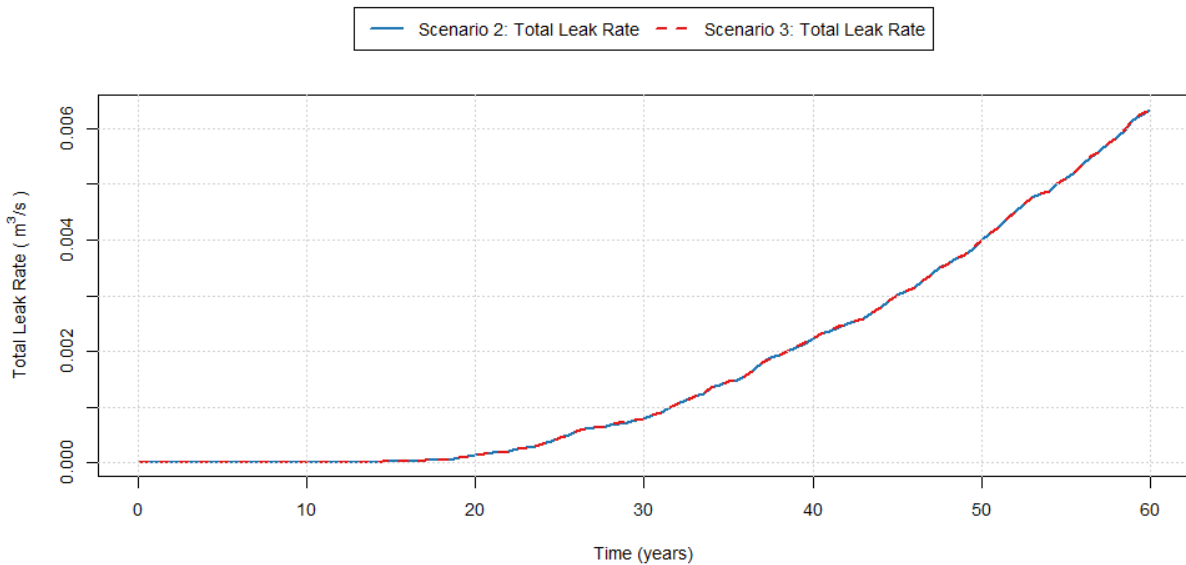


Figure 28: Mean total leak rate [m^3/s] for Scenario 2 (blue) and Scenario 3 (red).

Although this figure appears to show that these two leak rates are equivalent, inspection of the mean values shows that the total leak rate at 60 years for Scenario 3 is slightly higher than the total leak rate at 60 years for Scenario 2.

The occurrence of ruptures with ISI is expected to decrease for Scenario 3 when compared to Scenario 2. This is because when axial cracks exist, the probability of detection of both axial and circumferential cracks is greater than the probability of detection of circumferential cracks only. xLPR assumes that if a crack is detected the weld is fully repaired for the remainder of the plant life. Thus, Scenario 3 should experience fewer ruptures than Scenario 2. This result is confirmed in Figure 29.

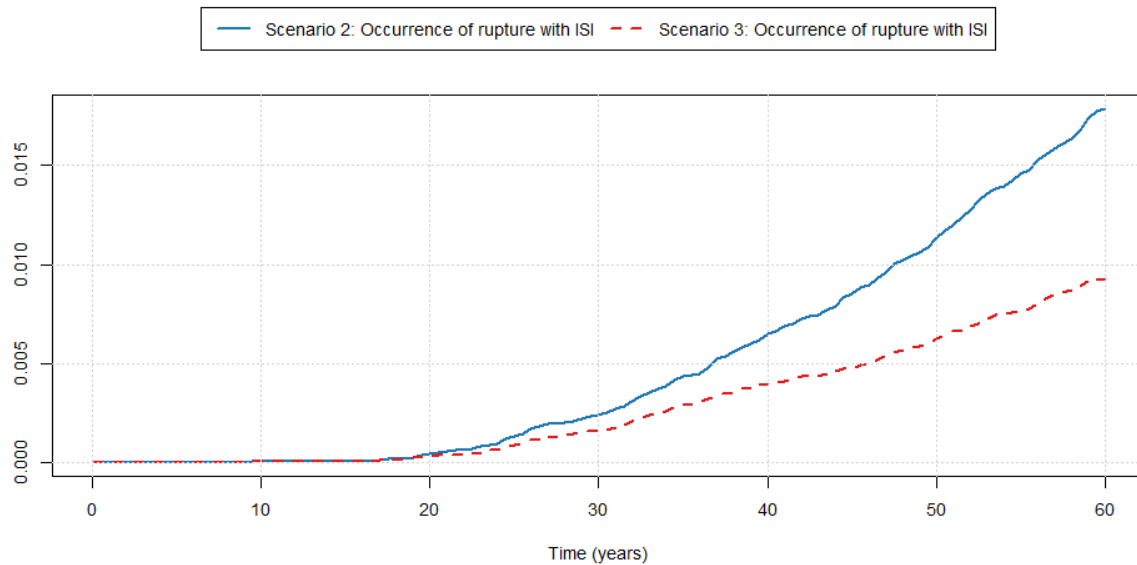


Figure 29: Mean probability of occurrence of rupture with ISI for Scenario 2 (blue) and Scenario 3 (red).

3.3 Scenario 3

3.3.1 Scenario Summary

The results and conclusions of a detailed analysis of sampling schemes and their impact on an understanding of the uncertainty found in the simulation results of Scenario 3 are given in this section.

Scenario 3 uses the following defining options in the xLPR code:

- PWSCC flaw initiation
- Circumferential and axial flaw orientation
- PWSCC flaw growth
- No mitigation or fatigue

This scenario forms the basis for Scenarios 4-9 which are subsequently defined by changes to mitigation options. Scenario 3 has a relatively higher probability of cracks, leaks, and ruptures.

The outputs of interest defined for this scenario for both circumferential and axial cracks are:

- Probability of occurrence of crack
- Probability of occurrence of leak
- Probability of occurrence of rupture
- Total Leak Rate
- Maximum crack depth
- Maximum crack inner half-length
- Maximum crack outer half-length

The comprehensive list of simulations that were used in the analysis of this scenario is shown below.

1. 100/50 (epistemic/aleatory sample size) with simple random sampling (SRS)
2. 500/50 with SRS
3. 500/50 with Latin hypercube sampling (LHS)
4. 500/50 with SRS and importance sampling (IS) on p2543 (multiplier on DM1 proportionality constant A) with target quantile of 0.95
5. 500/50 with LHS and IS on p2543 with target quantile of 0.95
6. 1000/50 with SRS
7. 1000/50 with LHS
8. 1000/50 with SRS and IS on p2543 with target quantile of 0.95
9. 1000/50 with LHS and IS on p2543 with target quantile of 0.95
10. 1000/50 with LHS and IS on p2543 with target quantile of 0.995
11. 10,000/100 with LHS

We assume a constant aleatory sample size of 50 for Runs 1-10. Preliminary analyses suggested that this aleatory sample size is sufficient for calculating a best-estimate of the probability of

occurrence. More discussion of aleatory sample size selection is provided in the advanced methods discussion presented in Section 4.1.

These sampling schemes were selected in order to gain insight into the performance of the xLPR code in relation to convergence of the best-estimate results for this scenario. The results of the analysis of these sampling schemes are summarized as follows. While many outputs of interest were defined for this scenario, the results of regression analyses indicate that the same variable, the multiplier on Direct Model 1 (DM1) proportionality constant A for the PWSCC crack initiation module, hereafter referred to as p2543, explained the most variability in all of the outputs. Because of this continuity and for the sake of brevity, the main result shown in this section is the probability of occurrence of circumferential ruptures. Importance sampling was applied to p2543 with the objective of improving the precision in outputs of interest, concentrating sampling on the 95th quantile. In Section 4.2, we examine how different quantile selections impacted convergence in the final results.

Overall, results confirm that increasing the epistemic sample size decreases the sampling uncertainty in the mean (best-estimate) of the output as expected, especially when moving from 100 to 500 epistemic samples. The decrease in sampling uncertainty in the mean of the output is also evident when transitioning from 500 to 1000 epistemic realizations and is likely attributable to the relatively high probability of crack and leak in this scenario. An epistemic sample size of 1000 appears to result in a stable (converged) ‘best estimate’ of the outputs for this scenario, but a sample of 500 with LHS does provide similar results for estimating the mean. The results of this scenario are more strongly impacted by the decision to use SRS or LHS. The use of LHS produces more stable results across sampling schemes without increasing computational time or analysis complexity. SRS results do not decrease the uncertainty in the mean of the results of interest as efficiently as LHS. The use of importance sampling was not found to have a strong impact on the results of this scenario, particularly at higher probabilities of occurrence. Importance sampling does result in ruptures occurring at earlier times, allowing estimation of the probability of rupture earlier, particularly when using SRS. In Section 4.2, we found that targeting the 0.999 quantile seemed to improve convergence more than targeting the 0.95 quantile.

3.3.2 Analysis Progression

An iterative approach was taken to select new sampling schemes for testing based on the comparative analysis of code runs. In general, the complexity of sampling options increased with each subsequent run. Each sampling scheme added to the list given in Section 3.3.1 provided new information regarding the impact of various sampling options on result convergence. Sampling scheme decisions were made in an attempt to increase precision in the mean output of interest. Decisions regarding sample size and which variables to importance sample are of particular interest due to their expected impact on mean estimation precision. Precision is quantified using the width of the 95% sampling confidence interval for the mean output of interest. The analysis of this scenario shows that the same conclusions can be drawn using any of the outputs given in the executive summary. Because of this consistency, probability of occurrence of circumferential ruptures will be used to exemplify the results of this scenario.

Runs 1 and 2 were used to determine if the epistemic sample size should be increased to reduce sampling uncertainty in the results. Figure 30 plots the mean probability of occurrence of circumferential ruptures for Run 1 and Run 2, along with the 95% bootstrap confidence intervals

(CIs) that describe the sampling uncertainty associated with the mean. The means for these two results begin to diverge at about 20 years and differ more at the end of the 60-year simulation. The 95% sampling confidence intervals shown in Figure 30 for Run 2 are narrower than those for Run 1. This comparison is also visualized in Figure 31, where the uncertainty associated with the mean for Run 2 is considerably lower than Run 1. Additionally, the uncertainty for Run 1 continuously increases more sharply over the course of the simulation time relative to Run 2. This reduction in result uncertainty shows that there is a benefit to increasing the epistemic sample size from 100 to 500. Because of this, an epistemic sample size of 500 was used to compare additional sampling options in subsequent runs.

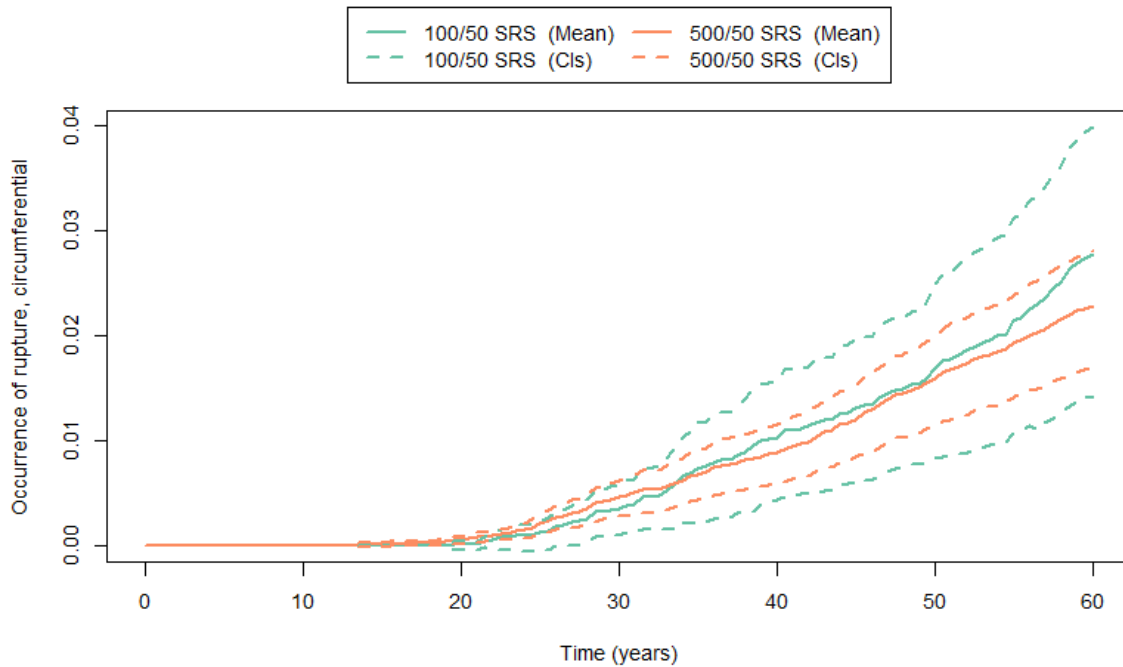


Figure 30: Mean probability of occurrence of circumferential ruptures (solid line) and 95% confidence intervals (dashed lines) for Scenario 3, Run 1 (green) and Run 2 (orange).

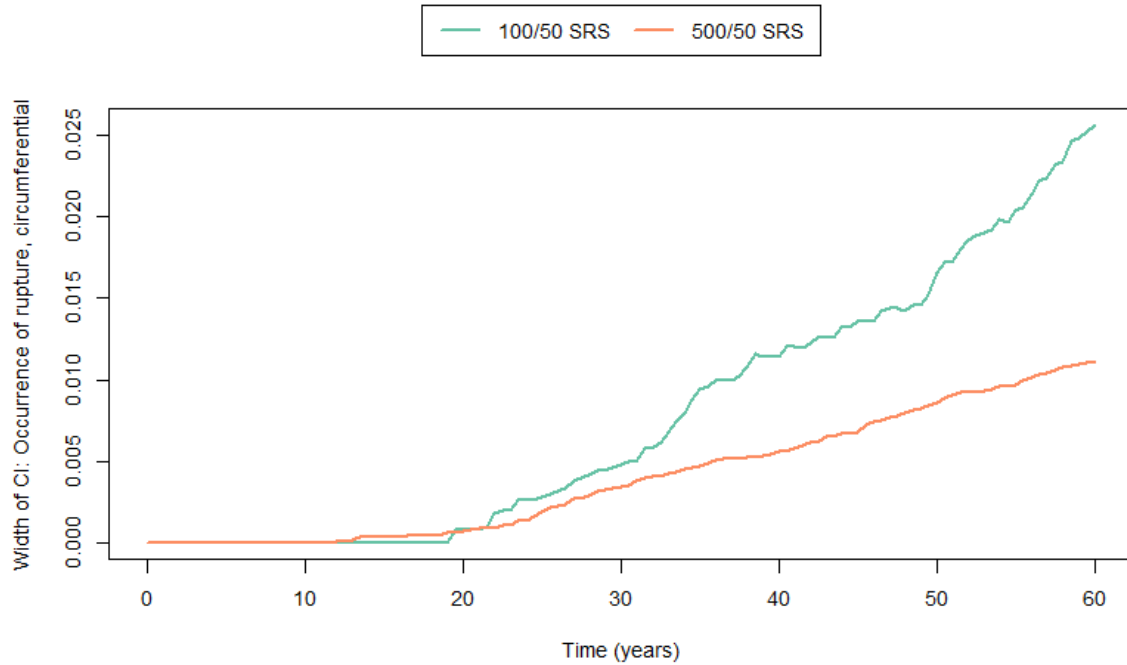


Figure 31: Width of 95% confidence interval for probability of occurrence of circumferential ruptures for Scenario 3, Run 1 (green) and Run 2 (orange).

The results of preliminary runs (Runs 1 through 3) identify variable p2543, the multiplier on DM1 proportionality constant A, as the most important variable for all of the scenario outputs. This importance can be seen both by a visual inspection of the relationship between this variable and the outputs of interest and through the use of a rank regression analysis. Figure 32 shows the scatter plots for this variable and the probability of occurrence of circumferential ruptures for Scenario 3, Runs 1 through 3.

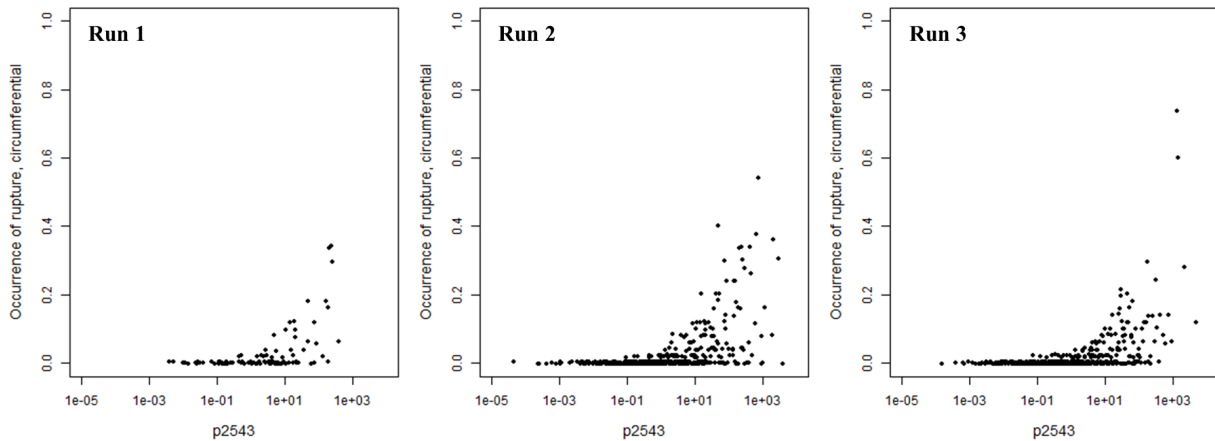


Figure 32: Scatter plots for variable p2543 (multiplier on DM1 proportionality constant A) (plotted on a log scale) and the probability of occurrence for circumferential rupture for Scenario 3, Run 1 (left), Run 2 (center) and Run 3 (right).

The relationship between p2543 and the occurrence of circumferential ruptures is not as strong as the relationship between this variable and other outputs including the occurrence of circumferential cracks, as demonstrated in Figure 33 and in the regression analysis results shown in Table 4.

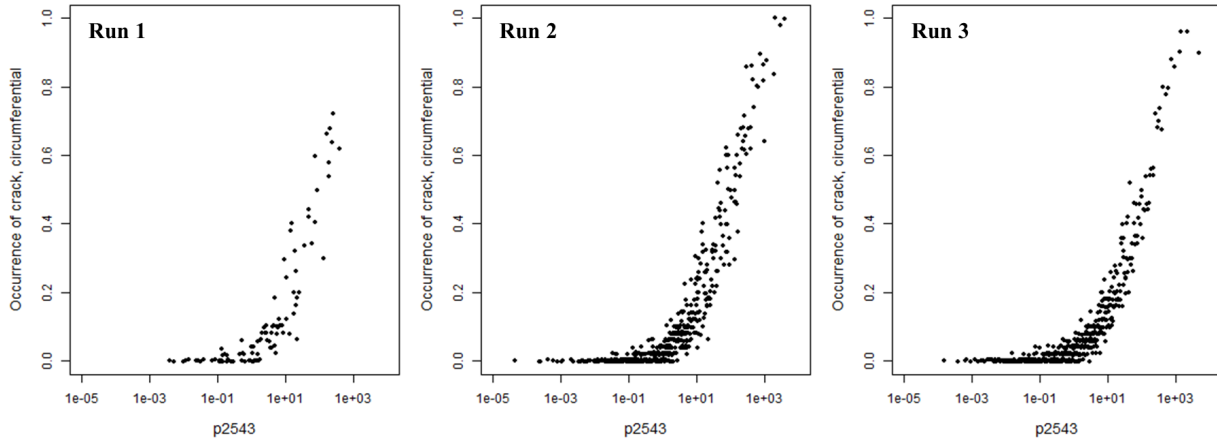


Figure 33: Scatter plots for variable p2543 (multiplier on DM1 proportionality constant A) (plotted on a log scale) and the probability of occurrence for circumferential cracks for Scenario 3, Run 1 (left), Run 2 (center) and Run 3 (right).

For the purpose of comparison, the scatter plots for two other variables, selected to show a lack of correlation with occurrence of circumferential ruptures, are provided in Figure 34 for Run 3. Variable p4352 (ID Pre-Mitigation Axial WRS) is shown because it is the second most important variable for probability of occurrence of circumferential rupture (see Table 4) while p2502 (Weld material ultimate strength) was selected for comparison due to its complete lack of importance to the final results. The lack of importance of p2502 is expected under the application of the DM1 PWSCC initiation model, making this variable useful for the illustration of a lack of correlation with the outputs of interest. These scatter plots demonstrate the strength of the impact that p2543 has on the output of interest.

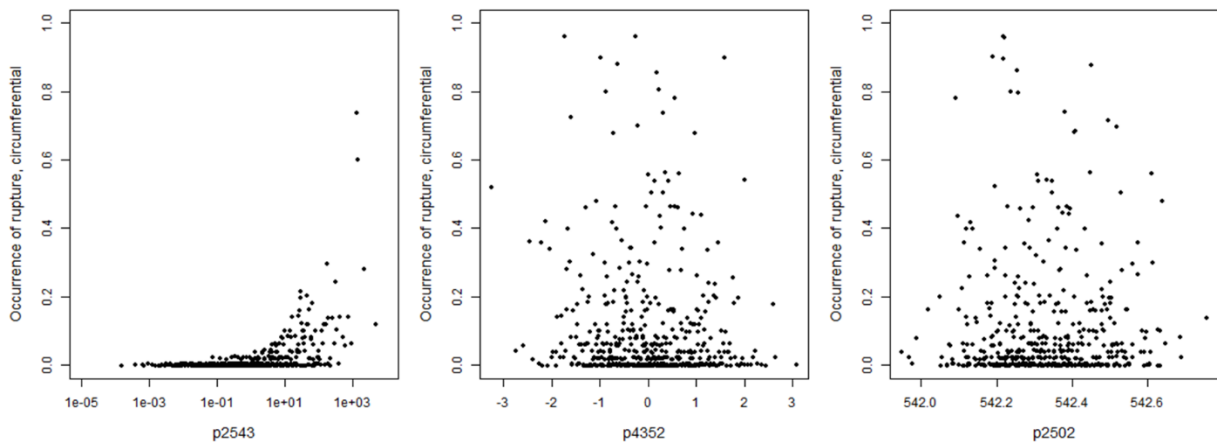


Figure 34: Scenario 3, Run 3 scatter plots for variables p2543 (left) (plotted on a log scale), p4352 (center), and p2502 (right).

The scatterplots for these variables show that p2543 has a definitive relationship with the output of interest as higher values of this variable correlate to higher occurrences of axial cracks. In contrast, there is not such a clear relationship between the sampled values of p4352 and p2502 and the probability of occurrence of circumferential ruptures.

A rank regression analysis was used to confirm the importance of variable p2453 to all of the outputs of interest. The results of this rank regression analysis, shown in Table 4, compare the standardized rank regression coefficient (SRRC) for the five most significant variables found in Run 3 for each output. The probability of occurrence of axial rupture is not shown in this table because there were not any axial ruptures observed for this scenario. Table 4 indicates that variable p2543 is the most important variable because the magnitude of its SRRC is much higher than the SRRCs of all of the other sampled variables across all of the outputs that are being analyzed for this scenario. The second most important variable changed based on the output under analysis.

Table 4: Summary of the stepwise rank regression analysis results for all outputs at year 60 for Scenario 3, Run 3.

Variable Identifier	Variable Name	SRRC	Variable Identifier	Variable Name	SRRC
Probability of Occurrence of Axial Cracks – $R^2 = 0.892$			Maximum Axial Crack Depth – $R^2 = 0.840$		
p2543	Multiplier proport. Const. A (DM1)	0.902	p2543	Multiplier proport. Const. A (DM1)	0.859
p4350	Hoop WRS Pre-mitigation	-0.228	p4350	Hoop WRS Pre-mitigation	-0.257
p5103	Log reg. intercept param., beta_0 (axial)	0.064	p2592	Comp-to-Comp Variab. Factor, fcomp	0.096
p5104	Log reg. slope param., beta_1 (axial)	0.050	p1102	Pipe Wall Thickness	-0.094
p1102	Pipe wall thickness	-0.039	p5103	Intercept, B0 (axial)	0.093
Probability of Occurrence of Circumferential Cracks – $R^2 = 0.847$			Maximum Circumferential Crack Depth – $R^2 = 0.808$		
p2543	Multiplier proport. Const. A (DM1)	0.923	p2543	Multiplier proport. Const. A (DM1)	0.873
p4352	Axial WRS Pre-mitigation	-0.106	p4352	Axial WRS Pre-mitigation	-0.191
p2505	Elastic Modulus, E	-0.049	p1102	Pipe Wall Thickness	-0.110
p9002	Surface Crack Dist Rule Modifier	0.037	p2592	Comp-to-Comp Variab. Factor, fcomp	0.088
p5101	Intercept, B0 (circ)	0.030	p2595	Charact. Width of Peak vs ECP, c	0.053
Probability of Axial Leak – $R^2 = 0.696$			Maximum Axial Inner Half-Length – $R^2 = 0.885$		
p2543	Multiplier proport. Const. A (DM1)	0.744	p2543	Multiplier proport. Const. A (DM1)	0.887
p4350	Hoop WRS Pre-mitigation	-0.230	p4350	Hoop WRS Pre-mitigation	-0.251
p2592	Comp-to-Comp Variab. Factor, fcomp	0.215	p5103	Intercept, B0 (axial)	0.072
p2595	Charact. Width of Peak vs ECP, c	0.162	p5104	Slope, B1 (axial)	0.059
p5103	Intercept, B0 (axial)	0.158	p1102	Pipe Wall Thickness	-0.042
Probability of Circumferential Leak – $R^2 = 0.528$			Maximum Circumferential Inner Half-Length – $R^2 = 0.788$		
p2543	Multiplier proport. Const. A (DM1)	0.601	p2543	Multiplier proport. Const. A (DM1)	0.856
p2592	Comp-to-Comp Variab. Factor, fcomp	0.245	p4352	Axial WRS Pre-mitigation	-0.204
p4352	Axial WRS Pre-mitigation	-0.243	p2592	Comp-to-Comp Variab. Factor, fcomp	0.115
p1102	Pipe Wall Thickness	-0.206	p1102	Pipe Wall Thickness	-0.112
p2595	Charact. Width of Peak vs ECP, c	0.162	p2595	Charact. Width of Peak vs ECP, c	0.090
Probability of Circumferential Rupture – $R^2 = 0.493$			Maximum Axial Outer Half-Length – $R^2 = 0.697$		
p2543	Multiplier proport. Const. A (DM1)	0.567	p2543	Multiplier proport. Const. A (DM1)	0.742
p4352	Axial WRS Pre-mitigation	-0.232	p4350	Hoop WRS Pre-mitigation	-0.230
p2592	Comp-to-Comp Variab. Factor, fcomp	0.232	p2592	Comp-to-Comp Variab. Factor, fcomp	0.217
p2595	Charact. Width of Peak vs ECP, c	0.213	p2595	Charact. Width of Peak vs ECP, c	0.162
p1102	Pipe Wall Thickness	-0.206	p5103	Intercept, B0 (axial)	0.156
Total Leak Rate – $R^2 = 0.703$			Maximum Circumferential Outer Half-Length – $R^2 = 0.526$		
p2543	Multiplier proport. Const. A (DM1)	0.744	p2543	Multiplier proport. Const. A (DM1)	0.598
p2592	Comp-to-Comp Variab Factor, fcomp	0.251	p2592	Comp-to-Comp Variab. Factor, fcomp	0.249
p1102	Pipe Wall Thickness	-0.189	p4352	Axial WRS Pre-mitigation	-0.247
p2595	Charact. Width of Peak vs ECP, c	0.182	p1102	Pipe Wall Thickness	-0.201
p4352	Axial WRS Pre-mitigation	-0.156	p2595	Charact. Width of Peak vs ECP, c	0.169

The importance of p2543 to all of the outputs of interest makes sense in the context of this scenario. This parameter is used in the Direct Model 1 PWSCC crack initiation model in xLPR, a key defining model for this scenario. Crack initiations are necessarily very important to occurrences of crack,

leak, and rupture, as well as to crack properties. The secondary importance of p4352 (inner diameter (ID) axial pre-mitigation weld residual stress (WRS)) for circumferential results and of p4350 (ID hoop pre-mitigation WRS) for axial results is also an expected result.

Although p2543 was found to be the most important variable related to the output of interest, the upper tail of the input distribution for p2543, which is the region most likely to result in cracks, is not well sampled using traditional SRS or LHS. Thus, importance sampling was applied to p2543 at the 95th quantile to oversample the upper tail in an attempt to generate more cracks. Figure 35 shows that this strategy did indeed generate more cracks as expected. This result confirms both the importance of variable p2543 to the results of this scenario and the advantages of the importance sampling capability in the xLPR code by showing that greater numbers of events can be observed for a fixed number of samples.

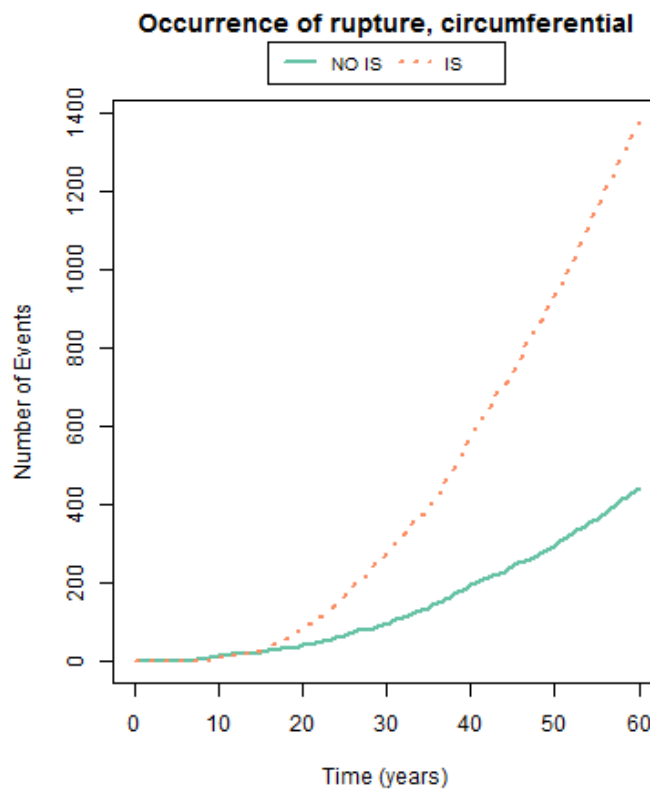


Figure 35: Number of circumferential rupture events occurring for Scenario 3, Run 3 (No IS) and Run 5 (IS with $q = 0.95$).

Figure 36 through Figure 39 show the effects of importance sampling on p2543 at the 95th percentile for a 500/50 sampling size under both SRS and LHS. For both the SRS and LHS sampling schemes, the uncertainty associated with mean output is not consistently reduced using importance sampling at this quantile level and with this sample size. However, using IS with the SRS 500/50 run did result in an estimated mean output that is more consistent with the larger sample size and LHS runs

than the SRS runs without IS. This result prompted the exploration of additional importance sampling quantiles (Run 10 at the 0.995 quantile).

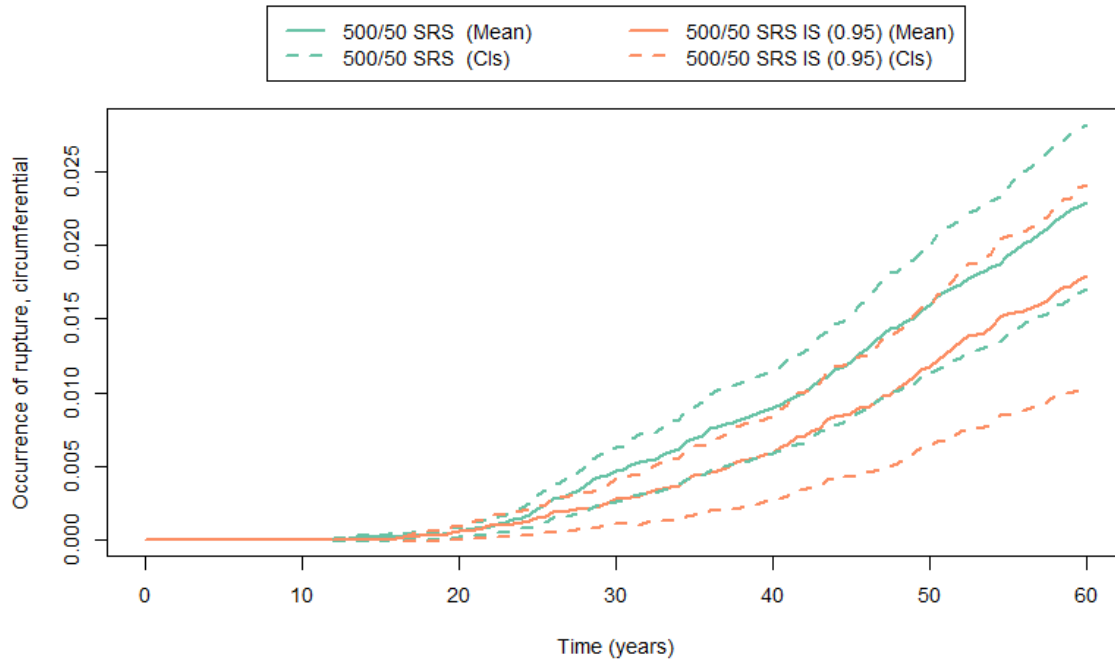


Figure 36: Mean probability of occurrence of circumferential ruptures (solid line) and 95% confidence intervals (dashed lines) for Scenario 3, Run 2 (green) and Run 4 (orange).

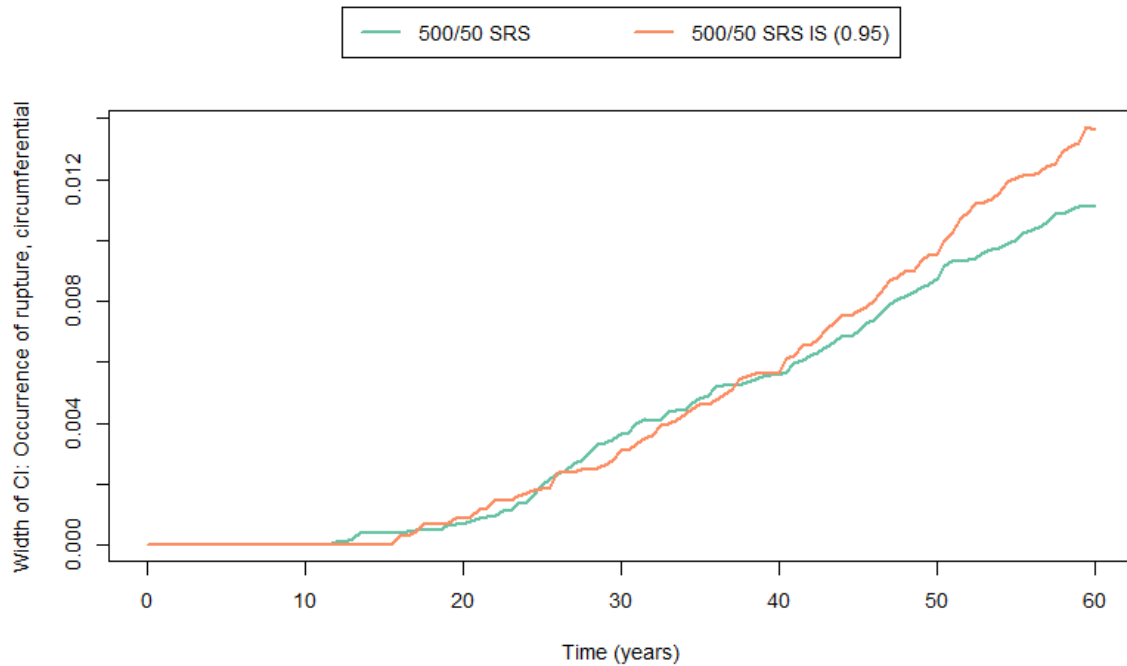


Figure 37: Width of 95% confidence interval for probability of occurrence of circumferential rupture for Scenario 3, Run 2 (green) and Run 4 (orange).

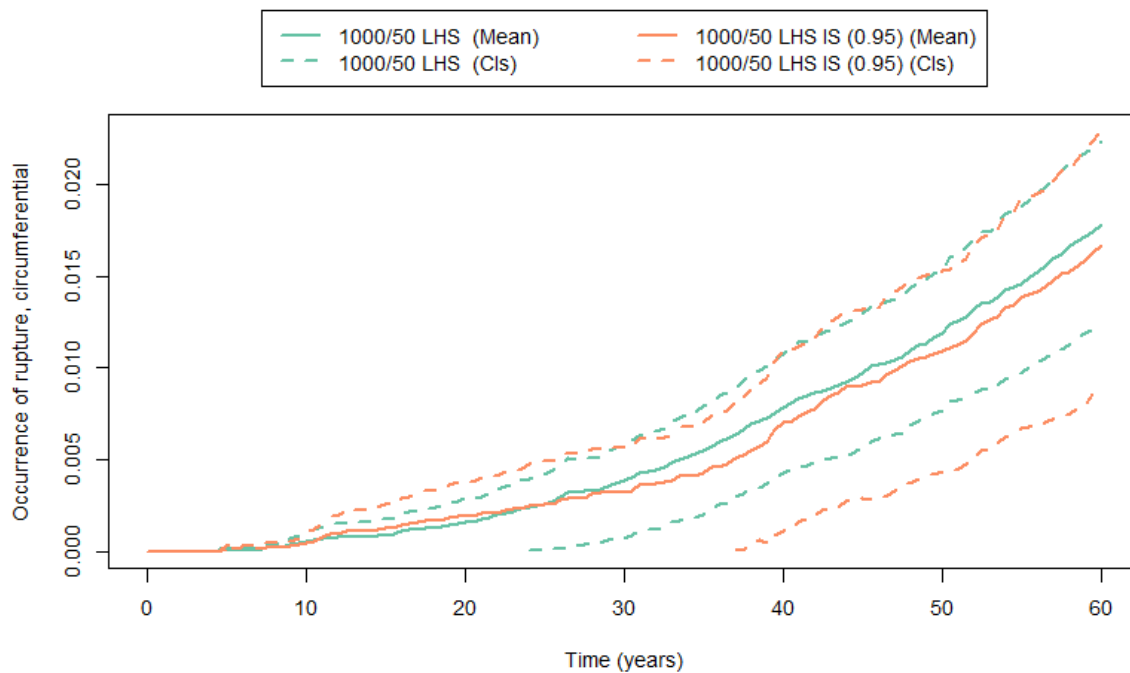


Figure 38: Mean probability of occurrence of circumferential ruptures (solid line) and 95% confidence intervals (dashed lines) for Scenario 3, Run 3 (green) and Run 5 (orange).

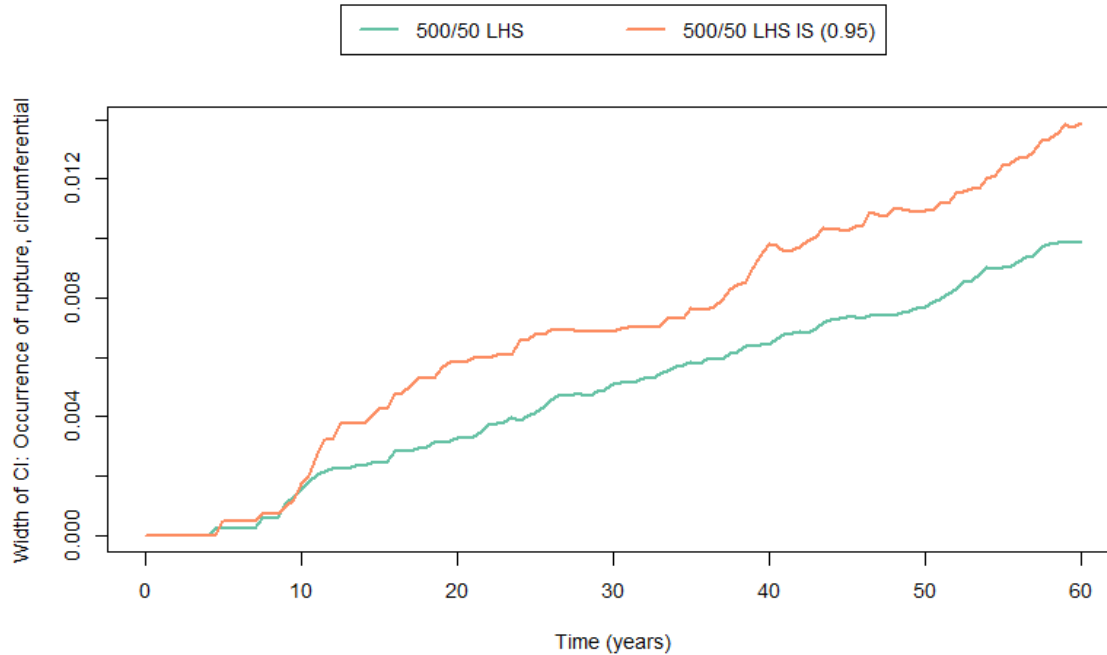


Figure 39: Width of 95% confidence interval for probability of occurrence of circumferential ruptures for Scenario 3, Run 3 (green) and Run 5 (orange).

Further studies on the impact of the epistemic sample size were also completed. It is possible that increasing the epistemic sample size from 500 to 1000 could further significantly reduce uncertainty, motivating the increase in sample sizes used for subsequent runs.

Figure 40 compares the results for epistemic sample sizes of 500 (Run 3) and 1000 (Run 7) using LHS and no importance sampling. The mean probability of occurrence of circumferential ruptures is similar through time for both of these runs. Figure 41 shows that there is a decrease in the width of the confidence interval about the mean at year 60 by about 0.004 when the epistemic sample size is increased from 500 to 1000; the relative efficiency of the increased sample size is even larger at earlier times (years ~10-30), when the probability of circumferential rupture is lower. Figure 31 shows that the width of the 95% CI decreased by about 0.015 at year 60 when the epistemic sample size was increased from 100 to 500. Based on this analysis, it can be seen that uncertainty in the mean does decrease with increasing sample size; the incremental decrease in uncertainty is relatively small at year 60 when moving from 500 to 1000 epistemic samples, but is larger at earlier time points when the failure probability is lower. The rate of event occurrence is relatively high for Scenario 3, meaning that it is expected that the confidence interval width will multiplicatively scale as a function of the square root of the number of epistemic samples (based on the Central Limit Theorem). The incremental gain in sample size may be expected to increase in other scenarios where events are rarer, as seen at earlier time points in Scenario 3.

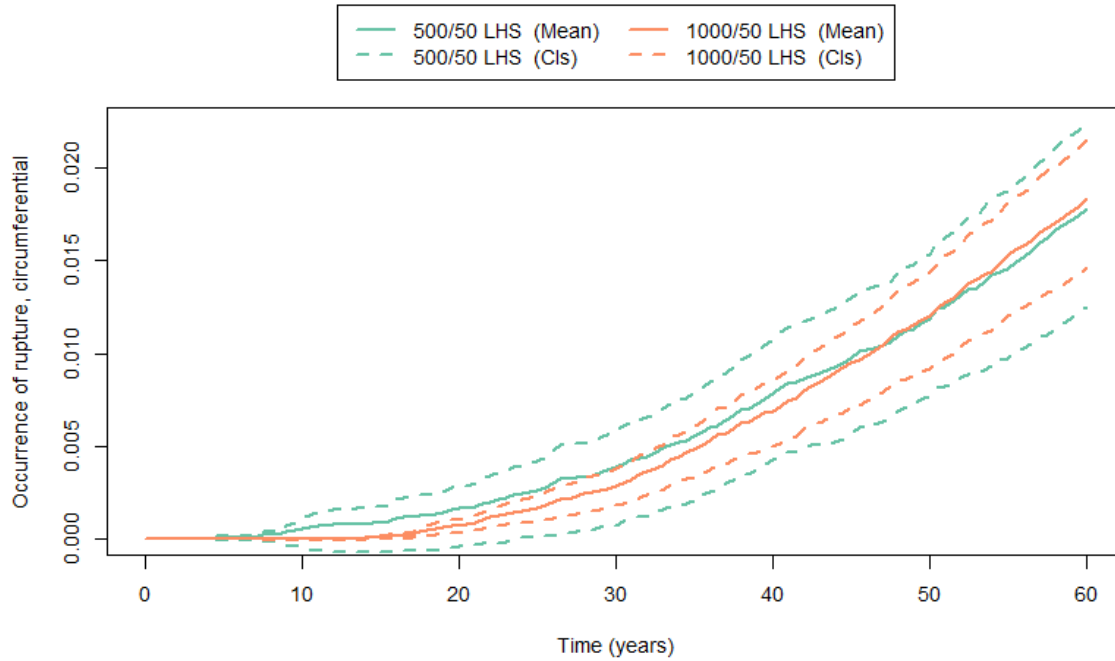


Figure 40: Mean probability of occurrence of circumferential ruptures (solid line) and 95% confidence intervals (dashed lines) for Scenario 3, Run 3 (green) and Run 7 (orange).

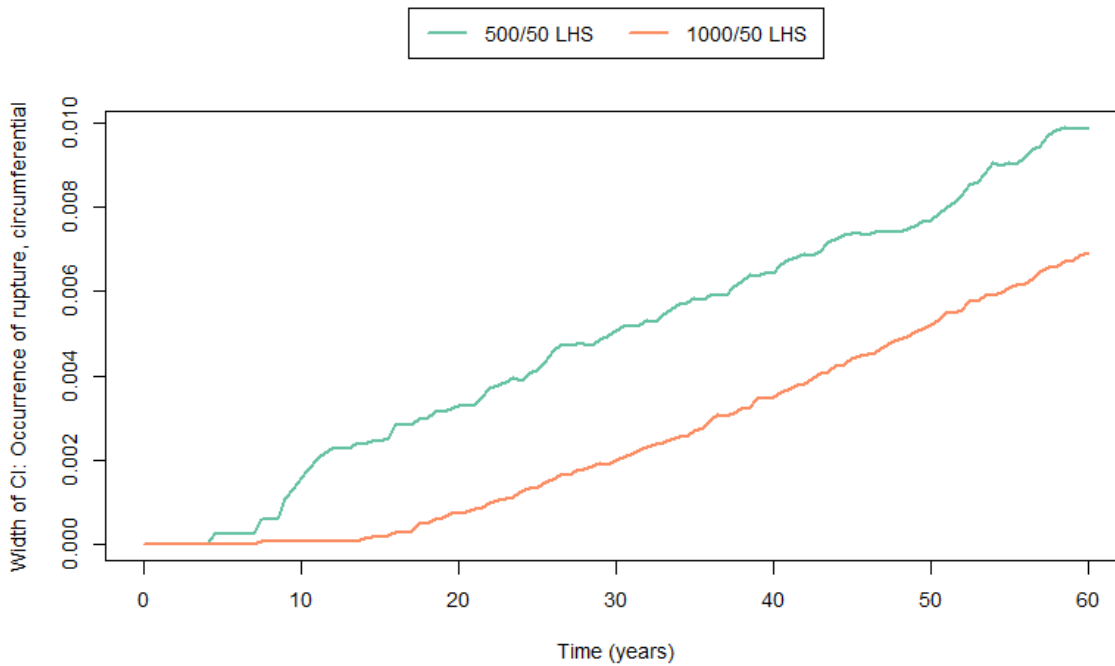


Figure 41: Width of 95% confidence interval for probability of occurrence of circumferential ruptures for Scenario 3, Run 3 (green) and Run 7 (orange).

Importance sampling was applied to this larger epistemic sample size to investigate its impact given these sampling selections. Figure 42 and Figure 43 compare sampling with and without IS when using LHS. Again, importance sampling p2543 at the 95th percentile for these particular runs is found to increase the width of the confidence intervals about the mean of the occurrence of circumferential ruptures in later time where the probability of rupture is higher. In early times (before 10 years) where the probability of rupture is lower, the application of importance sampling decreases the width of the confidence intervals, as shown in Figure 45. After 1 year, the confidence interval widths for occurrence of circumferential ruptures are larger under this importance sampling scheme (see Figure 47), further motivating additional Runs 10 and 11, as well as a study into the selected importance sampling quantile described in Section 4.2.

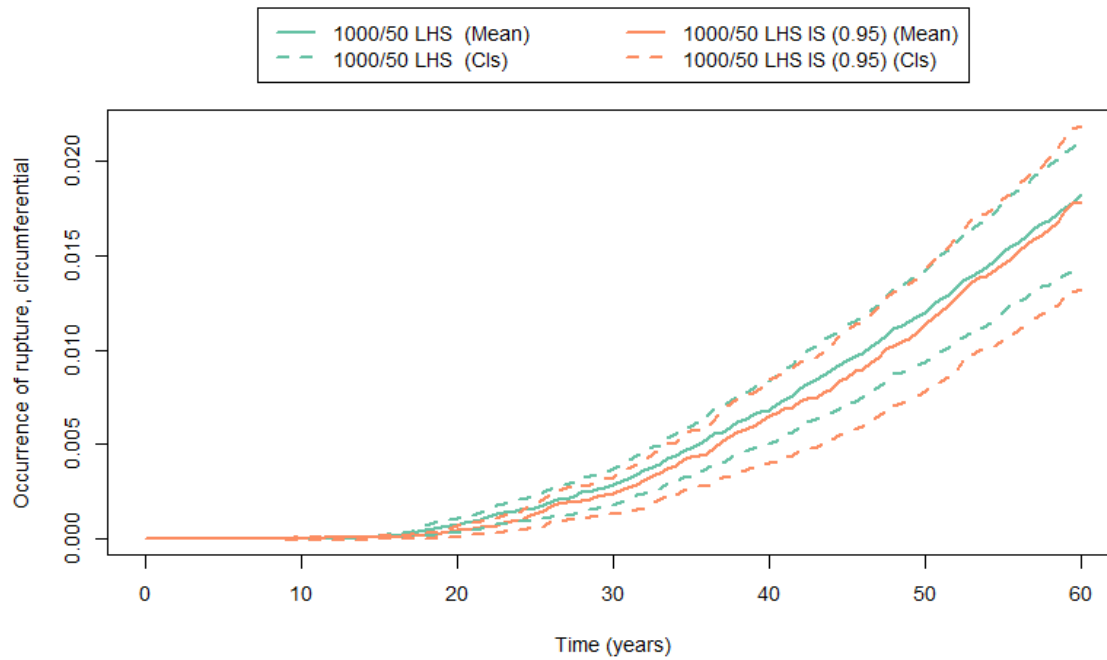


Figure 42: Mean probability of occurrence of circumferential ruptures (solid line) and 95% confidence intervals (dashed lines) for Scenario 3, Run 7 (green) and Run 9 (orange).

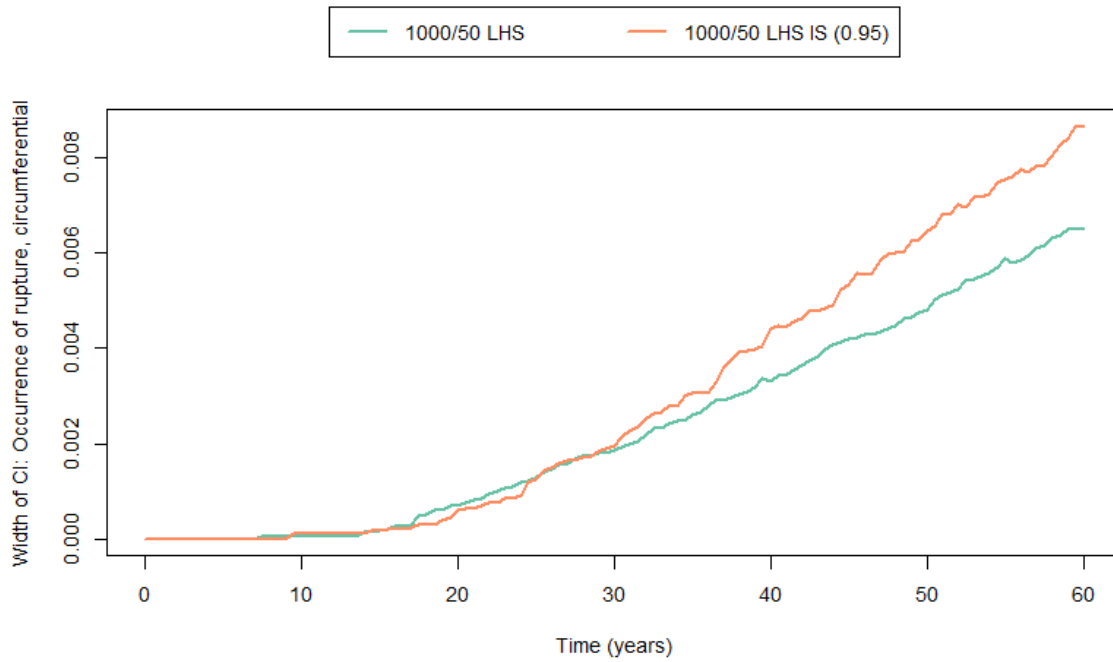


Figure 43: Width of 95% confidence interval for probability of occurrence of circumferential ruptures for Scenario 3, Run 7 (green) and Run 9 (orange).

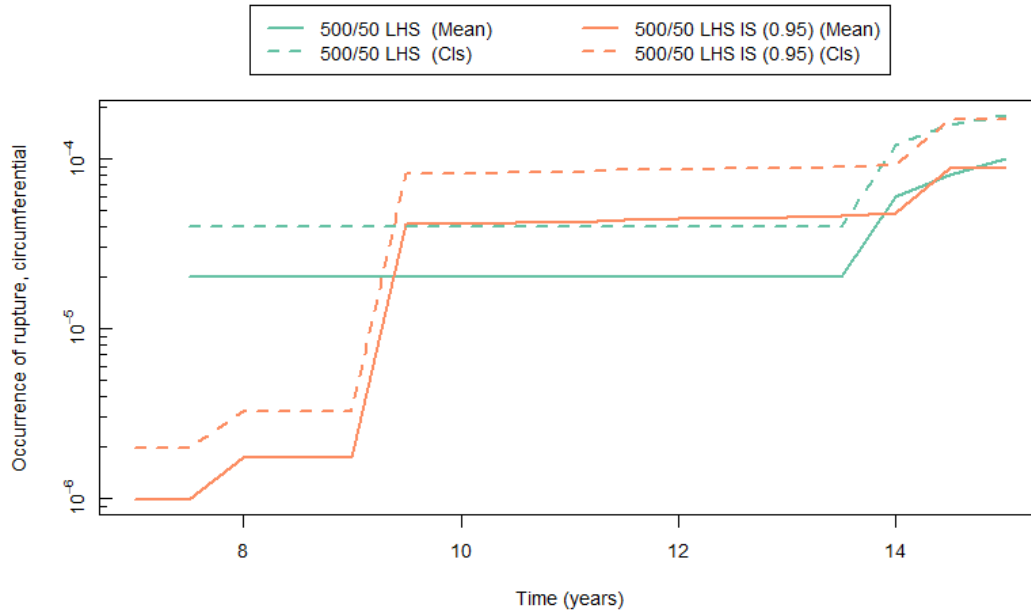


Figure 44: Mean probability of occurrence of circumferential ruptures (solid line) and 95% confidence intervals (dashed lines) for Scenario 3, Run 7 (green) and Run 9 (orange) up to 15 years with y-axis shown in log scale.

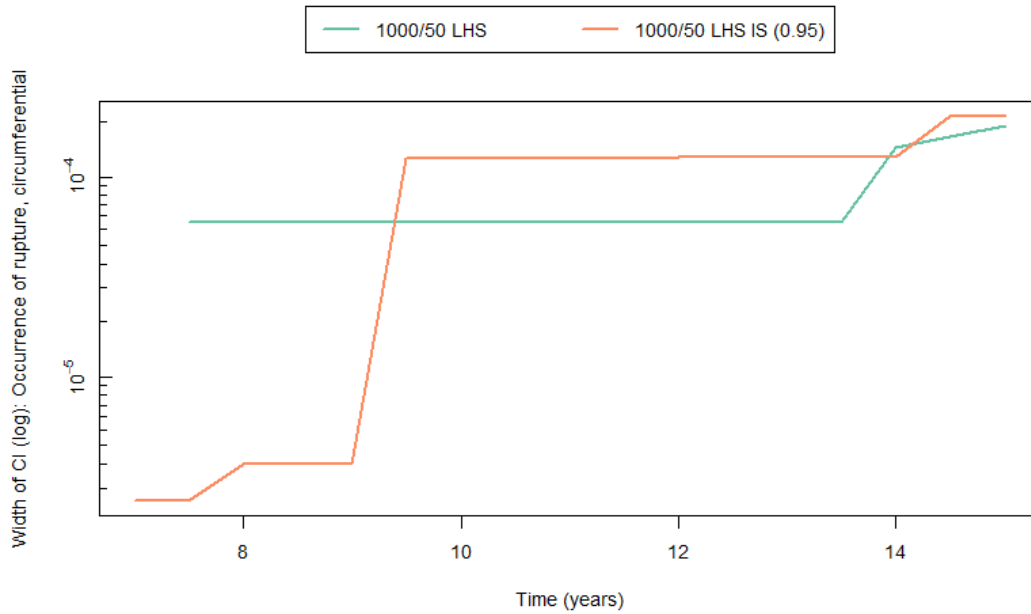


Figure 45: Width of 95% confidence interval for probability of occurrence of circumferential ruptures for Scenario 3, Run 7 (green) and Run 9 (orange) up to 15 years with y-axis shown in log scale.

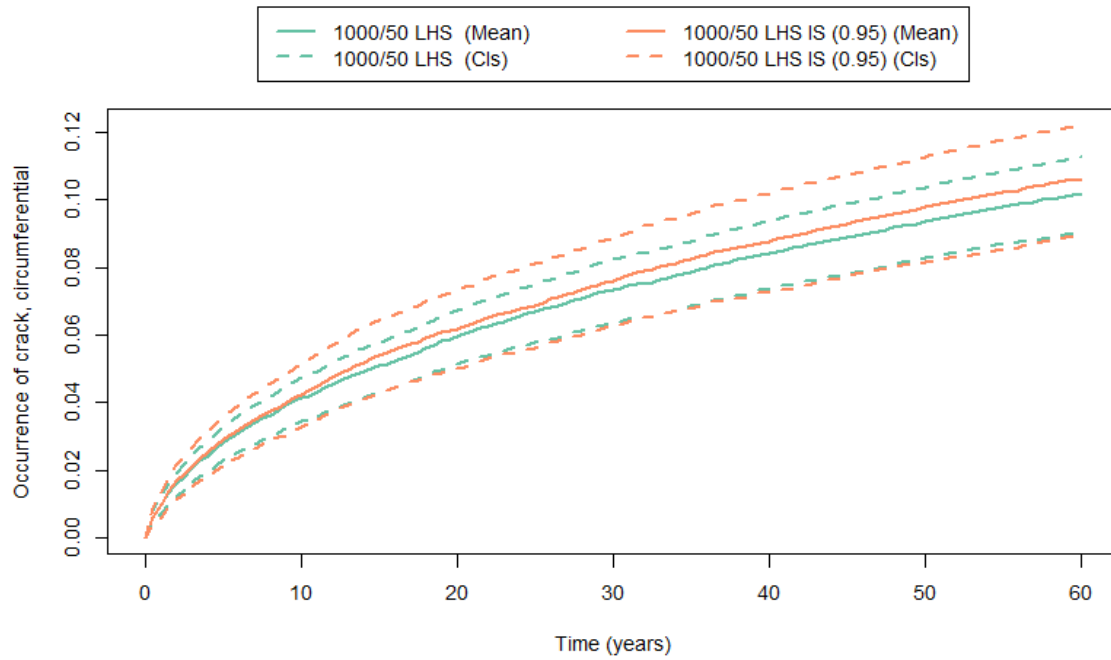


Figure 46: Mean probability of occurrence of circumferential cracks (solid line) and 95% confidence intervals (dashed lines) for Scenario 3, Run 7 (green) and Run 9 (orange).

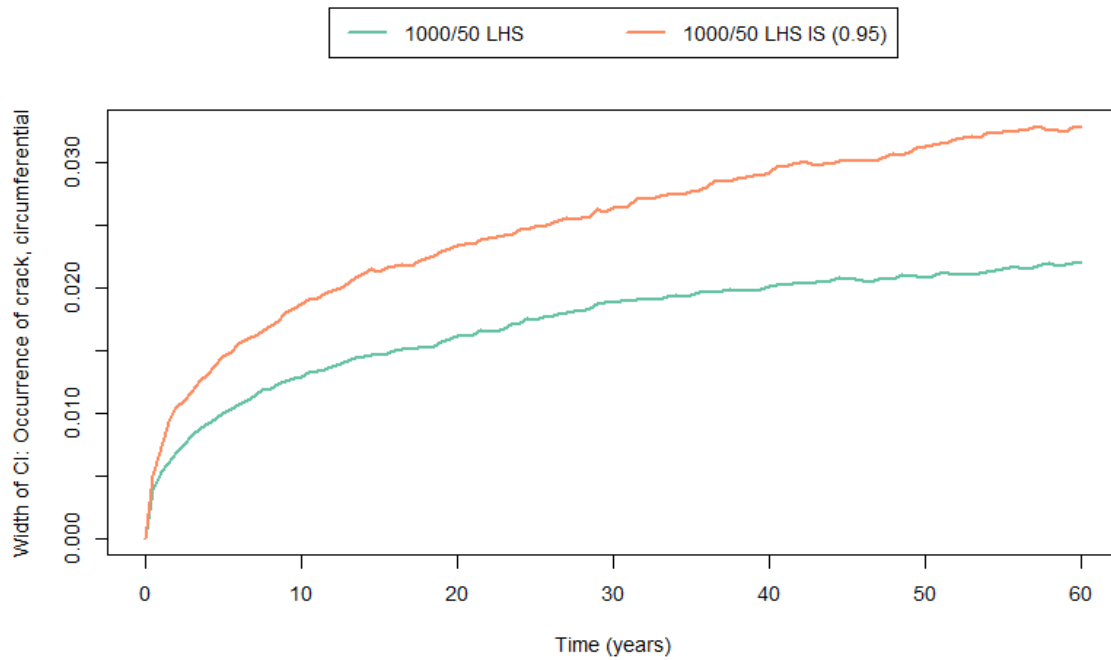


Figure 47: Width of 95% confidence interval for probability of occurrence of circumferential cracks for Scenario 3, Run 7 (green) and Run 9 (orange).

The current implementation of importance sampling in the xLPR code is dependent on the selected quantile about which importance samples will be concentrated. The selection of the 95th percentile for importance sampling on p2543 was found to increase the number of observed circumferential rupture events as expected. However, this sampling selection did not always decrease the uncertainty in the mean of the output of interest, occurrence of circumferential ruptures, or in additional outputs including the occurrence of circumferential cracks. This motivated the exploration of further quantiles for importance sampling on variable p2543.

A quantile value of 0.995 was selected for importance sampling of variable p2543 under a sampling scheme of 1000 epistemic realizations and LHS (Run 10). This value was based on a separate study discussed in Section 4.2 and in recommendations based on the results presented in (xLPR-SDD-FW, 2016). The width of the confidence intervals found under this importance sampling scheme are compared to the same run (Run 7) without importance sampling in Figure 49 below.

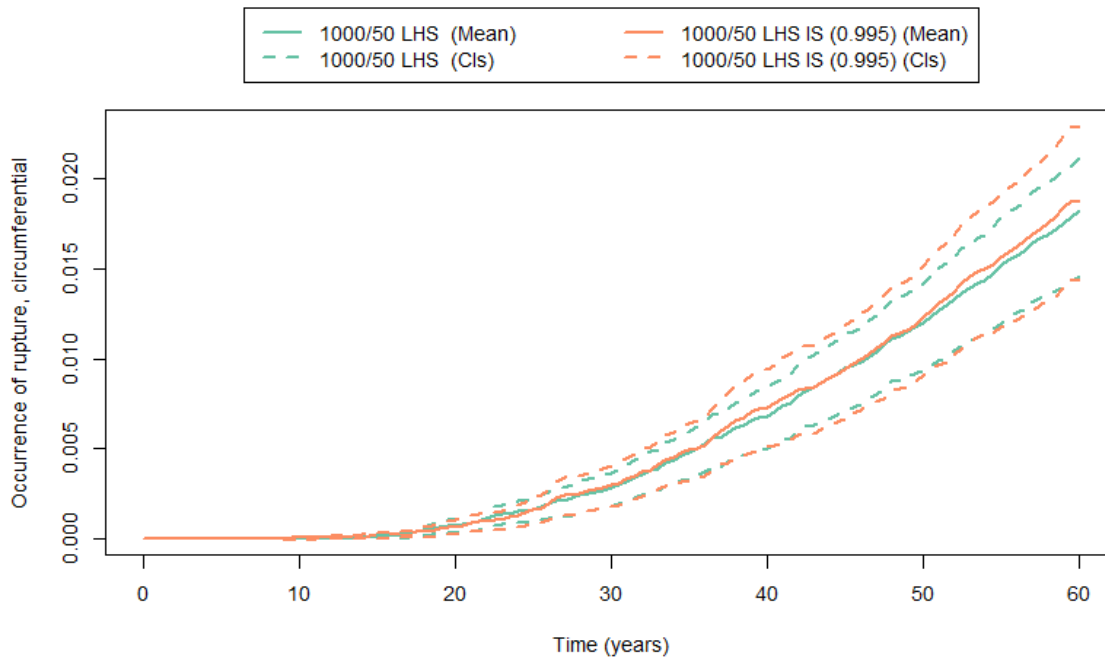


Figure 48: Mean probability of occurrence of circumferential ruptures (solid line) and 95% confidence intervals (dashed lines) for Scenario 3, Run 7 (green) and Run 10 (orange).

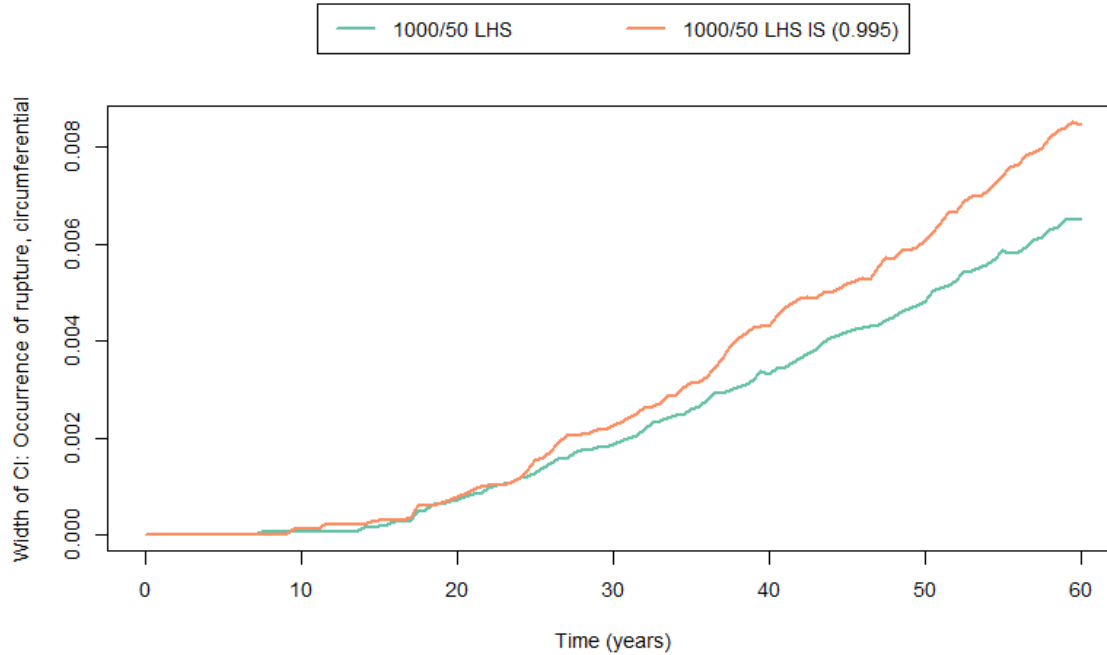


Figure 49: Width of 95% confidence interval for probability of occurrence of circumferential ruptures for Scenario 3, Run 7 (green) and Run 10 (orange).

The width of the confidence intervals about the outputs of interest were also compared to the results using a quantile value of 0.95. The widths of the confidence intervals calculated for the simulations using the two selected importance sampling quantile values for the occurrence of circumferential ruptures are shown in Figure 51. This figure shows that the new importance sampling quantile of 0.995 did not decrease the uncertainty in the mean result when compared to the quantile value of 0.95.

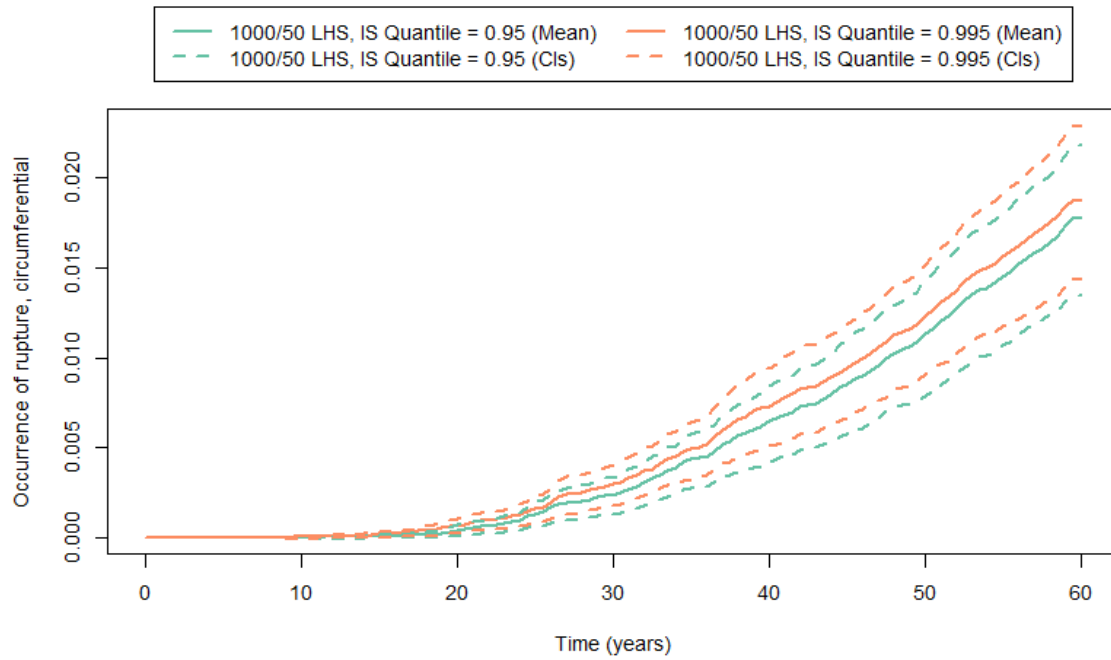


Figure 50: Mean probability of occurrence of circumferential ruptures (solid line) and 95% confidence intervals (dashed lines) for Scenario 3, Run 9 (green) and Run 10 (orange).

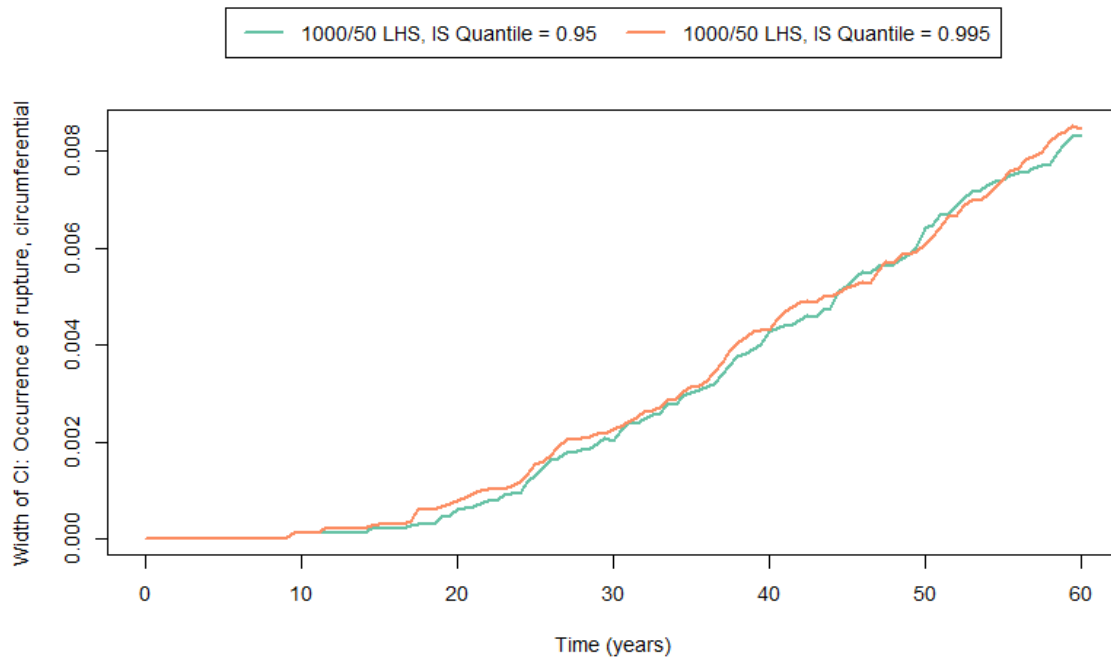


Figure 51: Width of 95% confidence interval for probability of occurrence of circumferential ruptures for Scenario 3, Run 9 (green) and Run 10 (orange).

The new quantile selection did have a stronger impact on the confidence interval width for occurrence of circumferential cracks. Figure 53 shows that the use of 0.995 as the importance sampling quantile for variable p2543 results in narrower confidence intervals than 0.95 for the occurrence of circumferential cracks. This is an expected result because variable p2543 is more strongly correlated with occurrence of circumferential cracks than it is with occurrence of circumferential ruptures (see Table 4).

While the selection of 0.995 as the importance sampling quantile for variable p2543 yields slightly better results for the occurrence of circumferential cracks than the same results using 0.95, the width of the confidence intervals are still larger at year 60 than those found without using importance sampling, as shown in Figure 55. However, the confidence intervals are slightly narrower in early time than when importance sampling is not applied, showing that some improvements are made in areas of lower probability as would be expected.

The selection of the ideal quantile for importance sampling on a single variable based on the final confidence interval width in areas of low probability is a pursuit that is convoluted by many factors. Conjoint parameter influences may motivate the selection of several variables for importance sampling. The importance distribution can have a large impact on efficiency gains (or losses) using IS. This importance sampling scheme and its dependence on the selected target quantile is described in Section 4.2. The complete investigation of these factors is outside of the scope of this report but could be an area of future work.

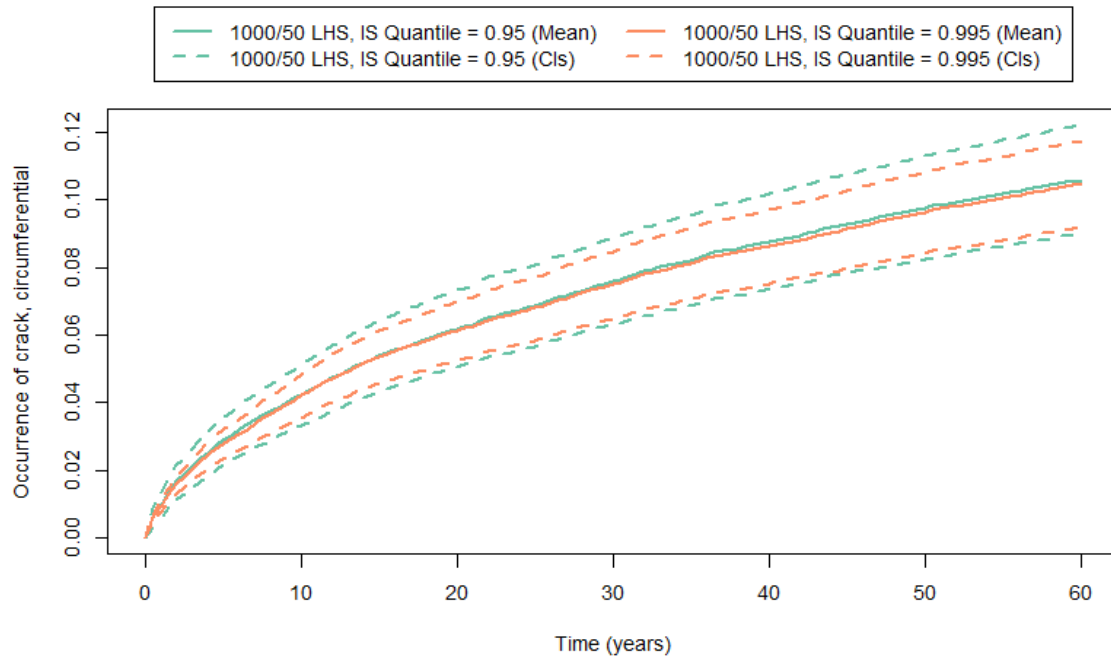


Figure 52: Mean probability of occurrence of circumferential cracks (solid line) and 95% confidence intervals (dashed lines) for Scenario 3, Run 9 (green) and Run 10 (orange).

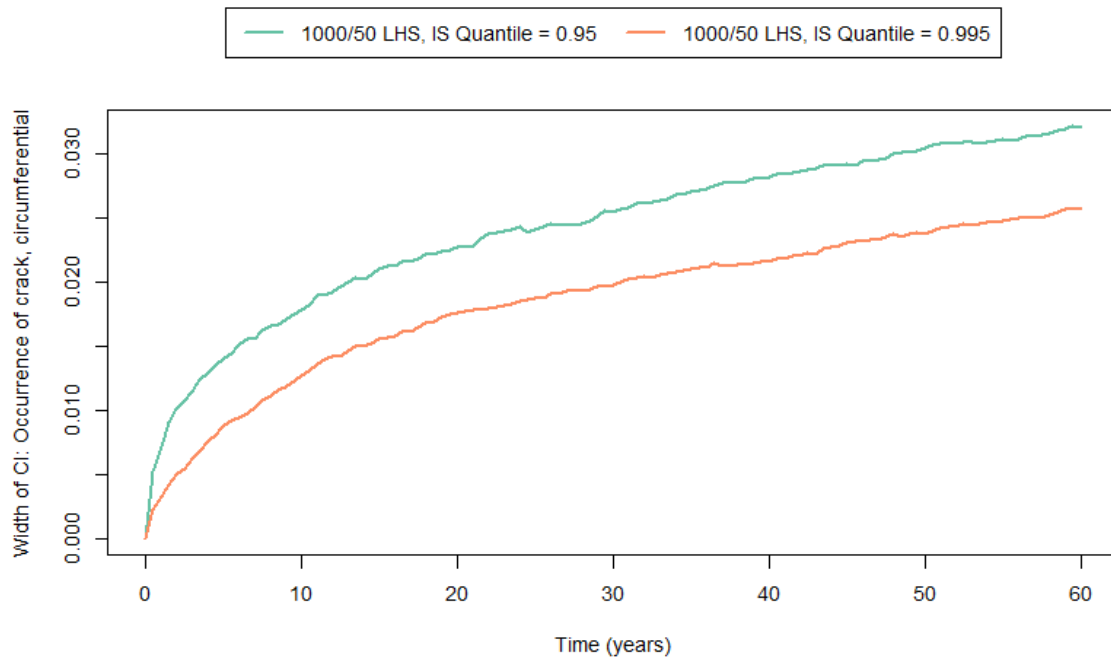


Figure 53: Width of 95% confidence interval for probability of occurrence of circumferential cracks for Scenario 3, Run 9 (green) and Run 10 (orange).

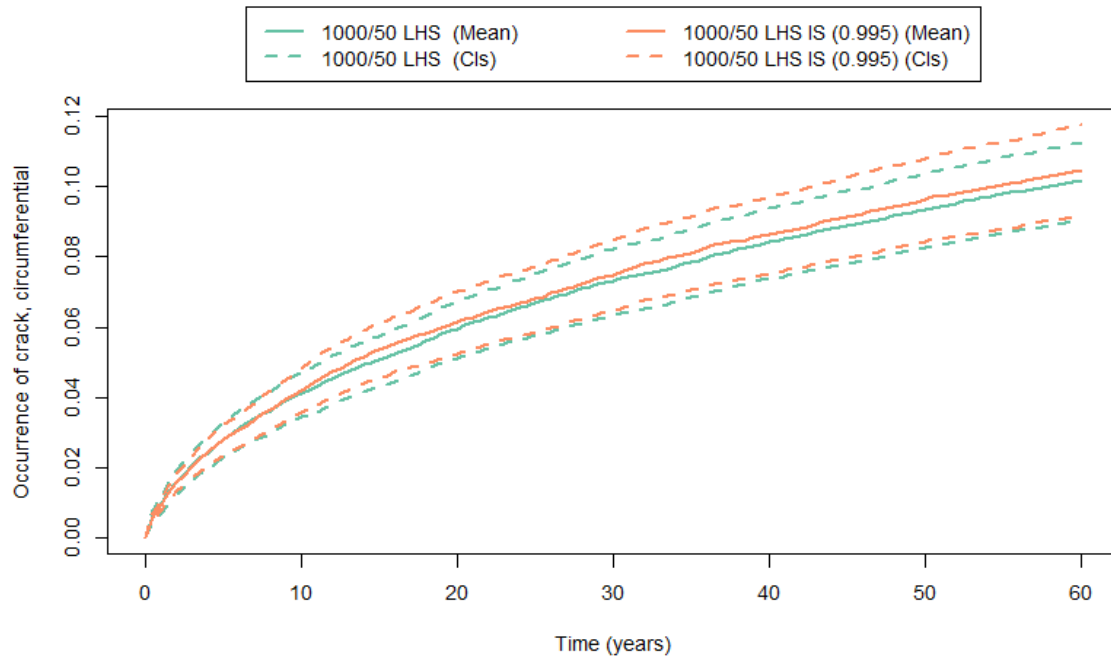


Figure 54: Mean probability of occurrence of circumferential cracks (solid line) and 95% confidence intervals (dashed lines) for Scenario 3, Run 7 (green) and Run 10 (orange).

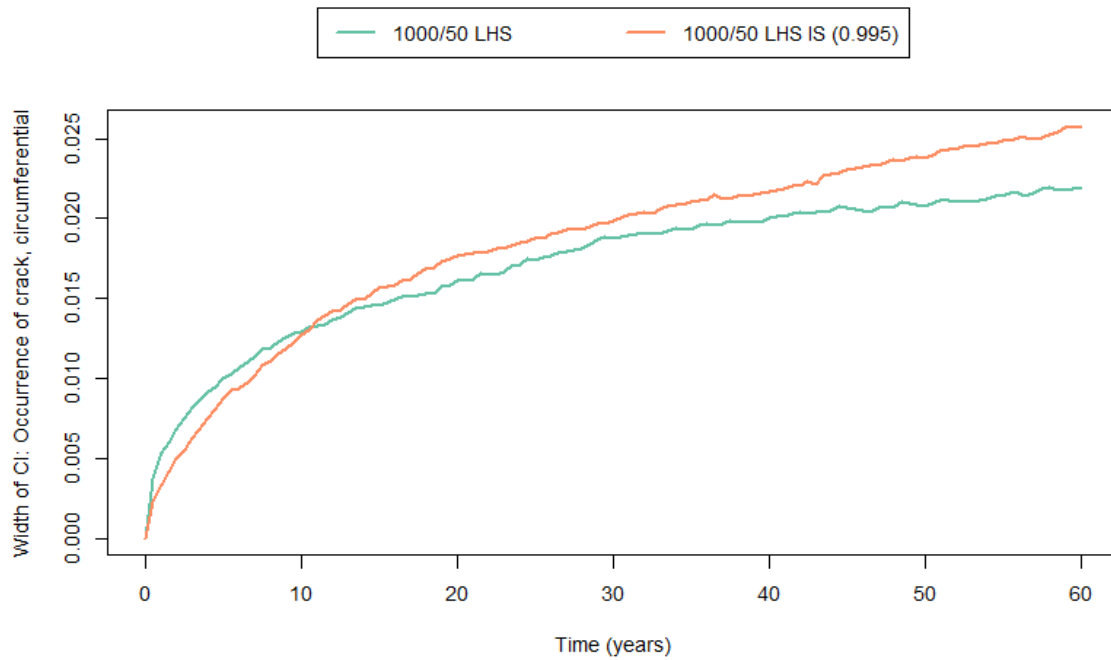


Figure 55: Width of 95% confidence interval for probability of occurrence of circumferential cracks for Scenario 3, Run 7 (green) and Run 10 (orange).

As a part of the investigation of the estimation of extremely low probability events such as the occurrence of ruptures under leak rate detection (described in more detail in Section 4.5), an additional run was added to the suite of Scenario 3 runs. Run 11 utilizes 10,000 epistemic realizations with 100 aleatory realizations each and LHS. In order to overcome memory limitations, this required the combination of 10 sets of runs with different random seeds and 1000 epistemic realizations with 100 aleatory realizations each. The results of this run are compared to Run 7 in Figure 56. The large epistemic sample size used in Run 11 significantly decreases the width of the confidence bounds about the mean of occurrence of circumferential ruptures, as expected. Estimates of the probability of circumferential rupture are similar between Runs 7 and 11.

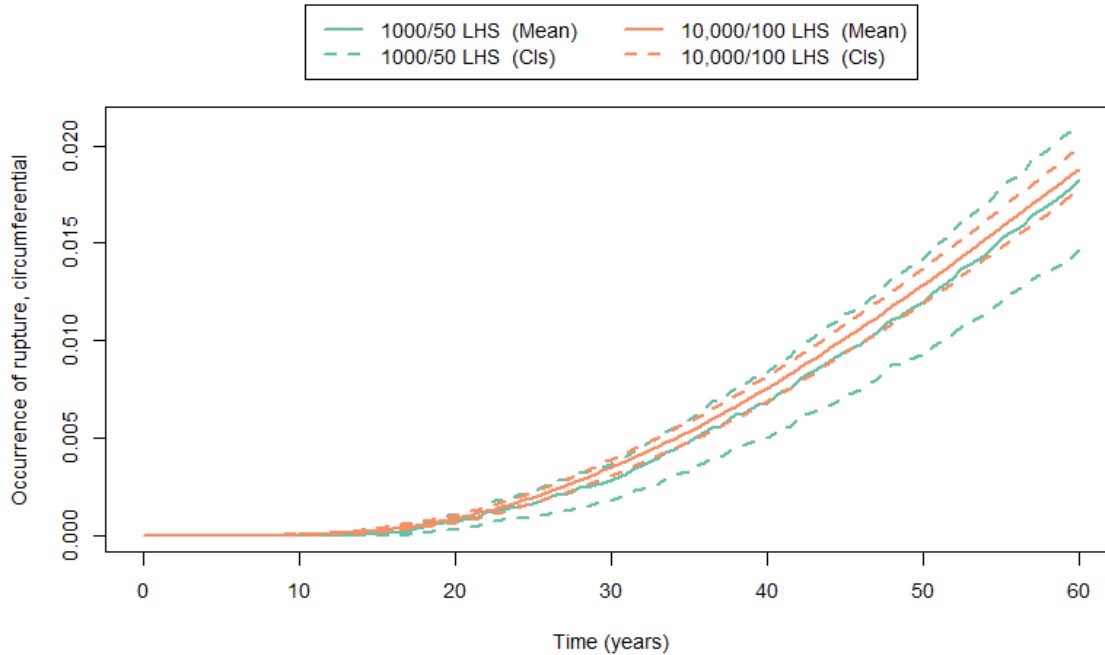


Figure 56: Width of 95% confidence interval for probability of occurrence of circumferential cracks for Scenario 3, Run 7 (green) and Run 11 (orange).

3.3.3 Overall Results

The iterative process and run-to-run comparisons described in the previous section led to the development of a collection of eleven runs uniquely distinguished by sampling option selections.

The examination of whether variations in sampling options lead to either converging or diverging results is common to many modeling analyses. Figure 57 shows the results for the mean probability of occurrence of circumferential cracks for Runs 1 through 10. This figure shows that runs using SRS with no IS, irrespective of sample size, have higher estimated event occurrences. A separate study into the dependence of these runs on the user-selected random seed was completed for Scenario 3 and is described in detail in Section 4.4. This study showed that the SRS runs with no IS shown in Figure 57 have mean results that are within sampling uncertainty; these runs were

found to have higher sampled values for p2543 on average than runs with the same sample size but different random seeds (by chance).

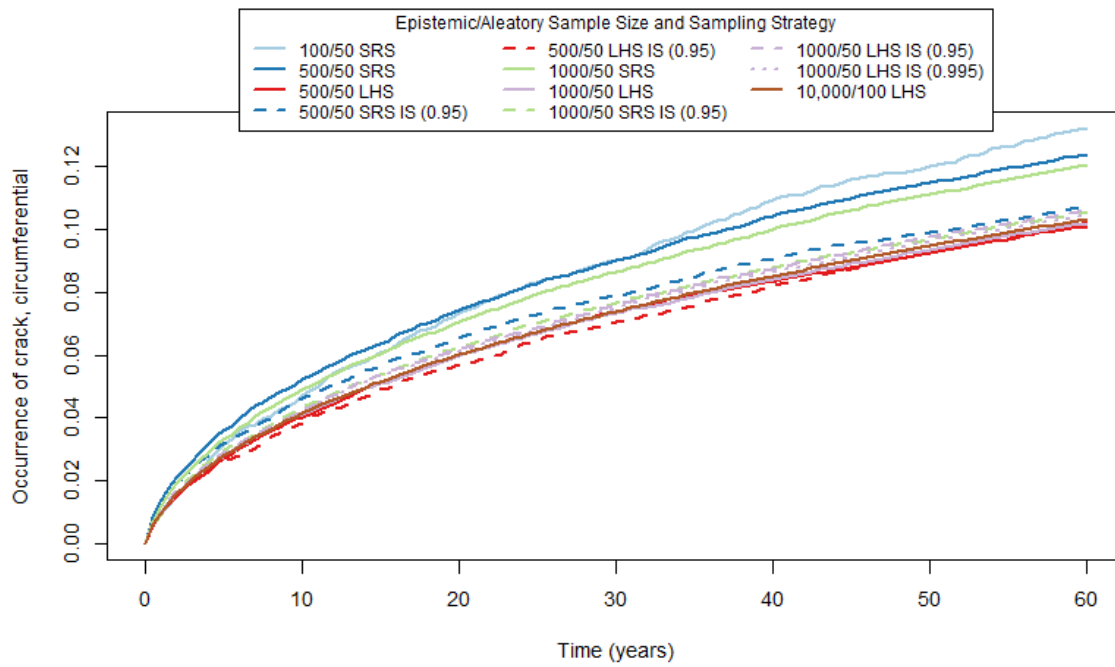


Figure 57: Mean probability of occurrence of circumferential cracks for Scenario 3, Runs 1 through 10.

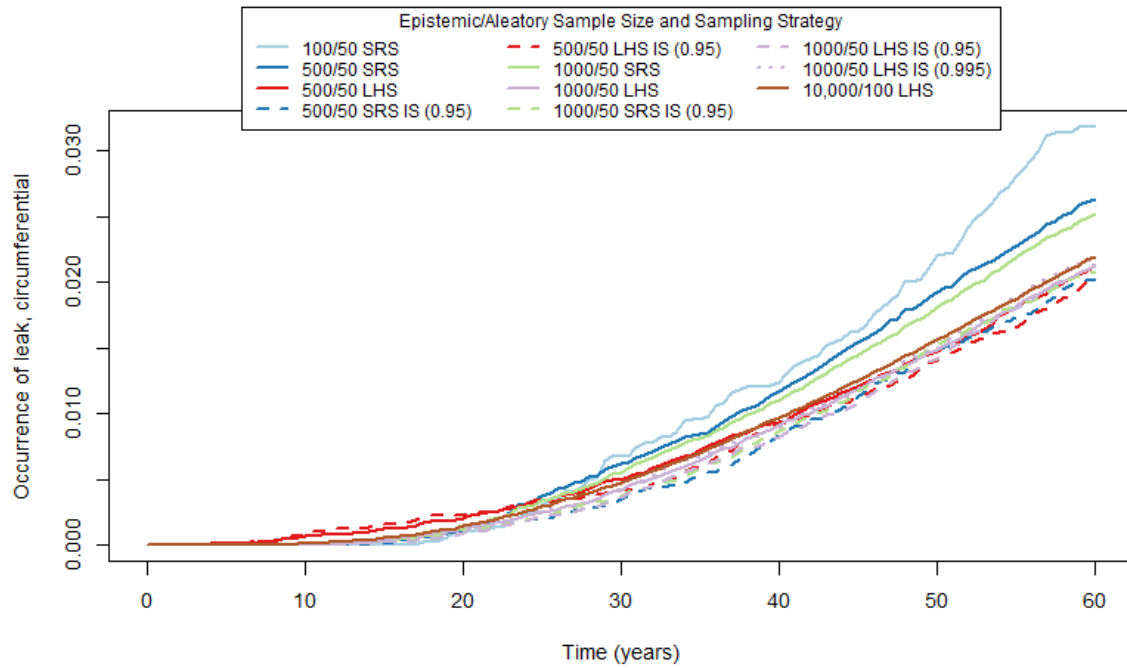


Figure 58: Mean probability of occurrence of circumferential leak for Scenario 3, Runs 1 through 10.

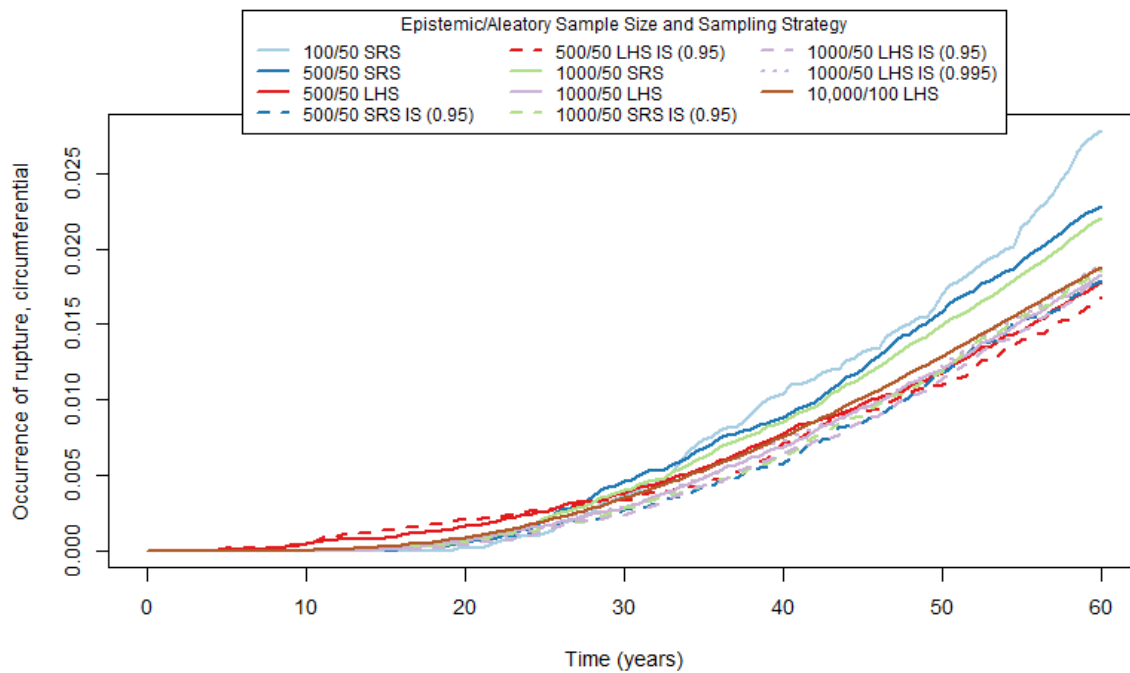


Figure 59: Mean probability of occurrence of circumferential rupture for Scenario 3, Runs 1 through 10.

The uncertainty associated with the result of interest, shown in this report in the form of 95% confidence intervals, can also be used to identify beneficial sampling selections for this scenario depending on the result of interest. Figure 31 and Figure 41 show how uncertainty is affected by the epistemic sample size. Choosing the final sample size will depend on both the computational time available and the level of uncertainty that the user is willing to accept. The uncertainty continuously decreases with increasing sample size, but the incremental gain from 500 to 1000 may not be worthwhile to a user who is concerned about computational time. Depending on user-needs pertaining to mean precision, smaller epistemic sample sizes may be sufficient.

In addition to the selection of an epistemic sample size, the choice of SRS or LHS may also impact the uncertainty in the result of interest. In Figure 57 and Figure 59, we see that LHS estimates of the probability of rupture were consistently more similar to Run 11, using the 10,000/100 sample size. In Section 4.2.2.1, we illustrate that time of first rupture was found to be earlier using LHS as compared to SRS, suggesting LHS may more thoroughly cover the input space. Because bootstrap confidence intervals cannot be accurately calculated for LHS, we explore this topic separately in Section 4.3 and compare estimates of sampling uncertainty in the mean using LHS versus SRS. Because LHS appear to provide efficiency gains and because computational time is not sacrificed when using LHS versus SRS, LHS sampling is a beneficial sampling selection.

Importance sampling at the 95th and 99.5th target quantiles does not necessarily impact the confidence interval width using these sampling options, as shown in Figure 43 and Figure 49. In Figure 57 and Figure 59, we see that SRS sampling without IS resulted in estimates that were rather different from the remaining estimates, while SRS sampling with IS estimates of the probability of rupture were consistently more similar to Run 11, using the 10,000/100 sample size, as well as the other LHS runs. Further, in Section 4.2.2.1, we illustrate that time of first rupture was found to be earlier using IS with SRS as compared to SRS only, suggesting IS allows better estimation of rare probabilities early in time. We expect importance sampling that the value of importance sampling will increase as the probability of event occurrence decreases.

Because the sampling scheme using LHS with 1000 epistemic, 50 aleatory samples with importance sampling on variable p2543 at the 0.95 quantile utilizes sampling options of interest and generated acceptable confidence intervals about the mean probabilities of interest, this sampling scheme is examined using additional analysis techniques in the subsequent sections.

3.3.4 Epistemic Uncertainty Analysis

An epistemic uncertainty analysis was conducted for Scenario 3, Run 9 (1000 epistemic and 50 aleatory samples and LHS with IS). The goal of the uncertainty analysis is to characterize epistemic uncertainty in the probabilities of occurrence.

Plots of the estimated probabilities of occurrence of circumferential crack, leak and ruptures are displayed in Figure 60 through Figure 62. These figures show the probability of occurrence for each epistemic realization (grey lines), the mean (red line), and the 5th, 50th, and 95th percentiles of the (blue, green, and purple lines, respectively). These percentiles correspond to estimated epistemic quantiles, which characterize uncertainty in probabilities of occurrence due to epistemic uncertainty. The lower and upper epistemic quantiles are constructed to encompass 90% of all the epistemic realizations in this run. These plots can be used to gain a better understanding of the spread of epistemic realizations and to analyze epistemic uncertainty intervals. For example, the

95th percentile for probability of circumferential ruptures is about 0.125 at year 60, meaning that 5% of the epistemic realizations have larger than a 12.5% chance of rupture by year 60. For leak, the 5th and 50th percentiles are uniformly zero up to year 60, indicating that more than 50% of the realizations did not experience a leak in the first 60 years.

The epistemic realizations for leak are slightly shifted to the right when compared to the same result for cracks. This intuitive result indicates that for this scenario, cracks are soon followed by leaks. This relationship is further illustrated in Figure 63, which displays the mean and 95th percentile for crack and leak. This plot also shows that the spread between the mean and the 95th percentile is larger for crack than it is for leak.

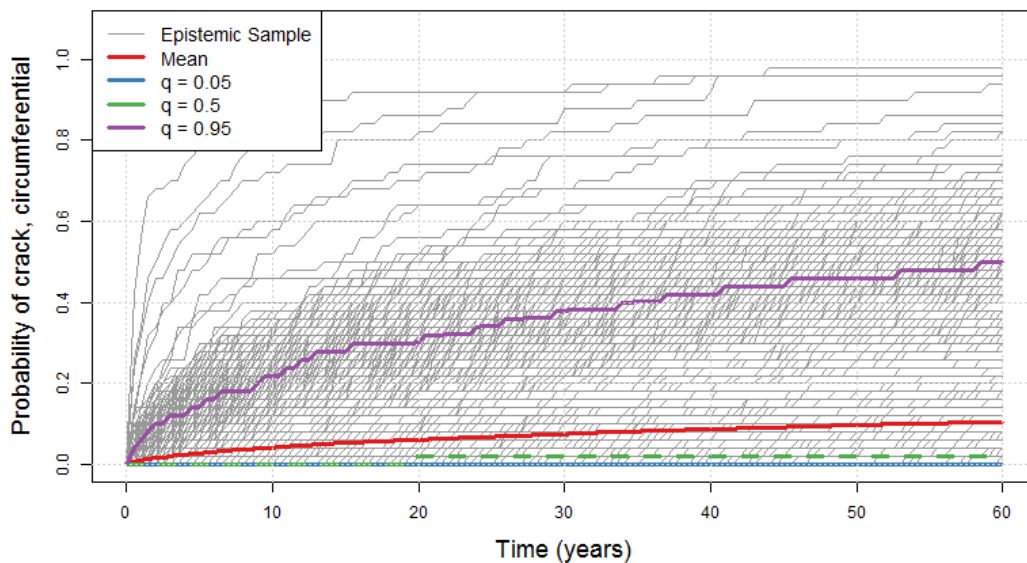


Figure 60: Probability of occurrence of circumferential crack for each epistemic realization (grey), the mean (red), and the 5th (blue), 50th (green), and 95th (purple) percentiles for Scenario 3, Run 9.

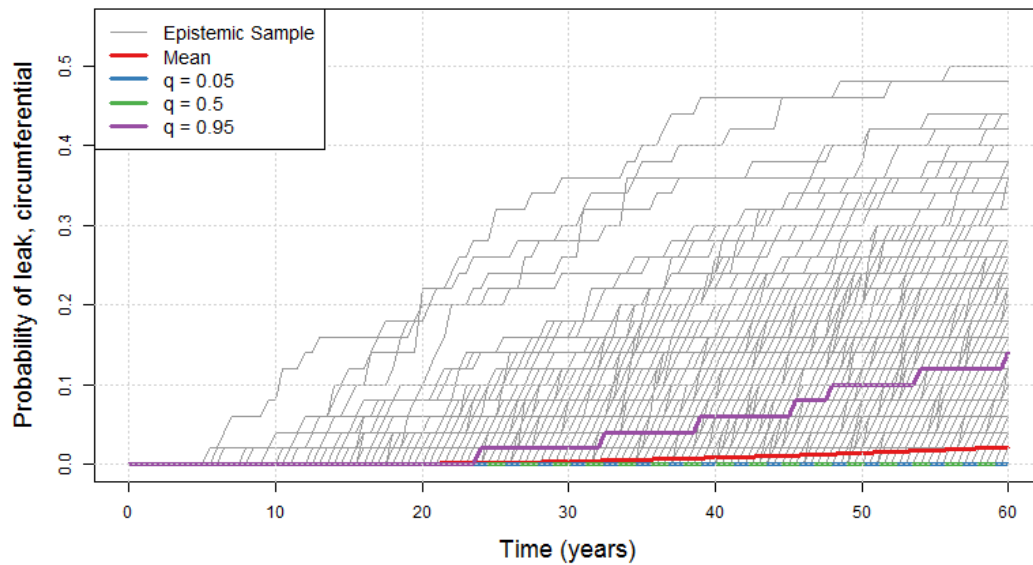


Figure 61: Probability of occurrence of circumferential leak for each epistemic realization (grey), the mean (red), and the 5th (blue), 50th (green), and 95th (purple) percentiles for Scenario 3, Run 9.

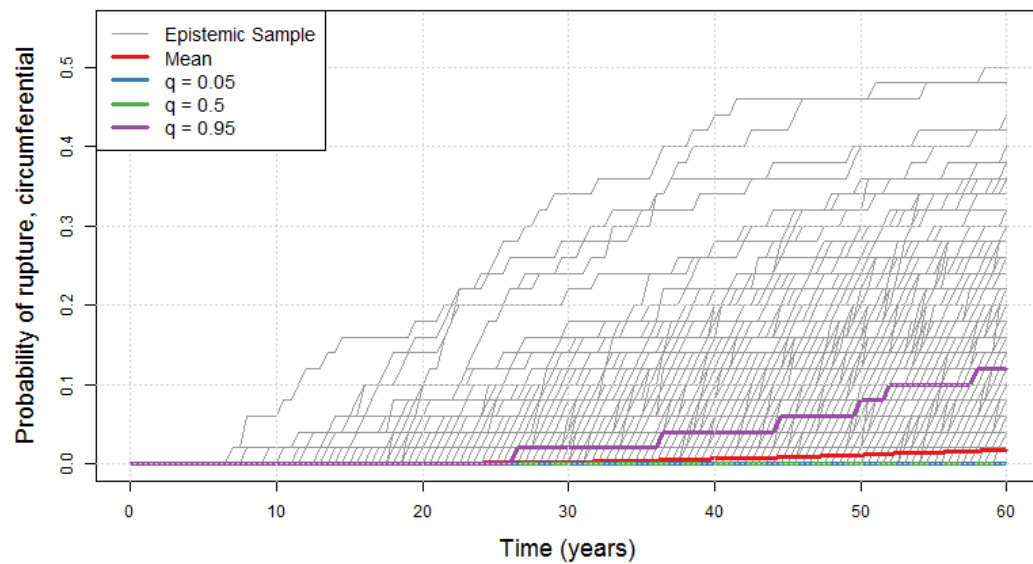


Figure 62: Probability of occurrence of circumferential rupture for each epistemic realization (grey), the mean (red), and the 5th (blue), 50th (green), and 95th (purple) percentiles for Scenario 3, Run 9.

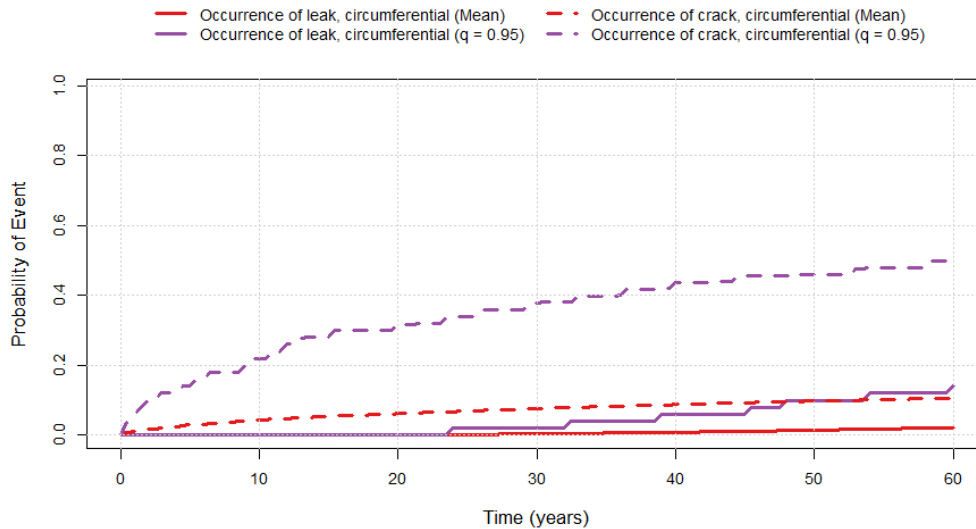


Figure 63: Mean (red) and 95th percentile (purple) for probability of occurrence of circumferential leak (solid) and crack (dashed) for Scenario 3, Run 9.

Recall that, in Section 3.3.2, 95% sampling confidence intervals were constructed for the mean. Analogous sampling confidence intervals could be placed on the percentile estimates to characterize sampling uncertainty in the epistemic quantiles, as shown in Figure 64. These bounds can be interpreted as: we have 95% confidence that the true 95th percentile of epistemic samples is within these bounds. Extreme quantiles are more difficult to estimate than the mean, so if the user is concerned with estimating the 95th percentile, a larger sample size will be necessary. Importance sampling of p2543 will likely improve precision in extreme percentile estimation. Selecting a sampling scheme for percentile estimation is beyond the scope of this report. Lastly, we do not account for aleatory sampling uncertainty, and only use an aleatory sample size of 50; in Section 4.1, we illustrate how aleatory sampling uncertainty can artificially inflate epistemic uncertainty.

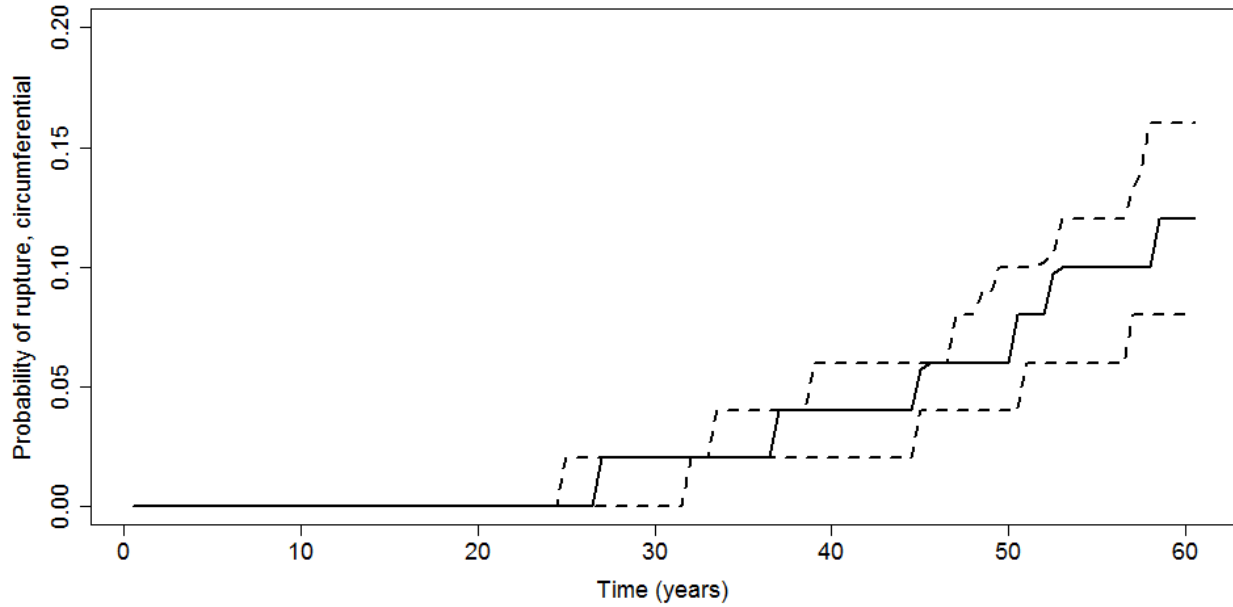


Figure 64: 95% sampling confidence intervals (dashed) on the 95th percentile of the epistemic realizations (solid) for the 1000 epistemic samples with LHS and IS.

3.3.5 Convergence Analysis

In order to gain better insight into the convergence of the sampling options used in Run 9(1000 epistemic, 50 aleatory samples with LHS and IS), this sampling scheme was re-run five times with different random number seeds for both the epistemic and aleatory loops. Each replicate was used to give an estimate of the mean probability of occurrence of circumferential ruptures. Figure 65 shows a comparison of these estimates. The replicates produce similar estimates with a magnitude of variation between them on the order of 10^{-2} at year 60. The five replicates were also used to estimate an overall mean value and a 95% prediction interval was created around this mean, as seen in Figure 66. The prediction interval represents the region that the next convergence run would be expected to fall within 95% percent of the time. The interval appears has a consistent and relatively narrow interval width after year 30. Using 1000 epistemic samples with LHS and IS on p2543 appears to produce stable estimates of the mean of occurrence of circumferential ruptures, regardless of the random seed used in the simulation.

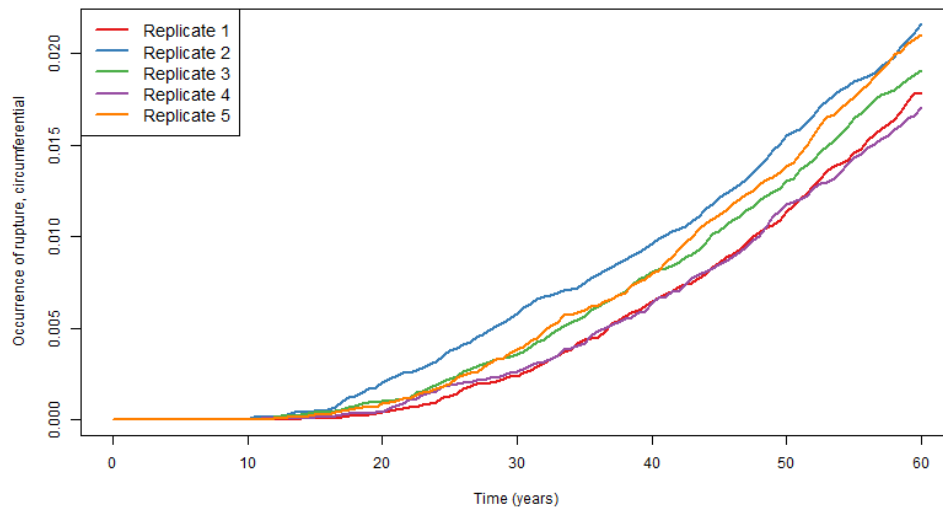


Figure 65: Comparison of estimated mean probability of occurrence of circumferential rupture for five convergence replicates for Scenario 3.

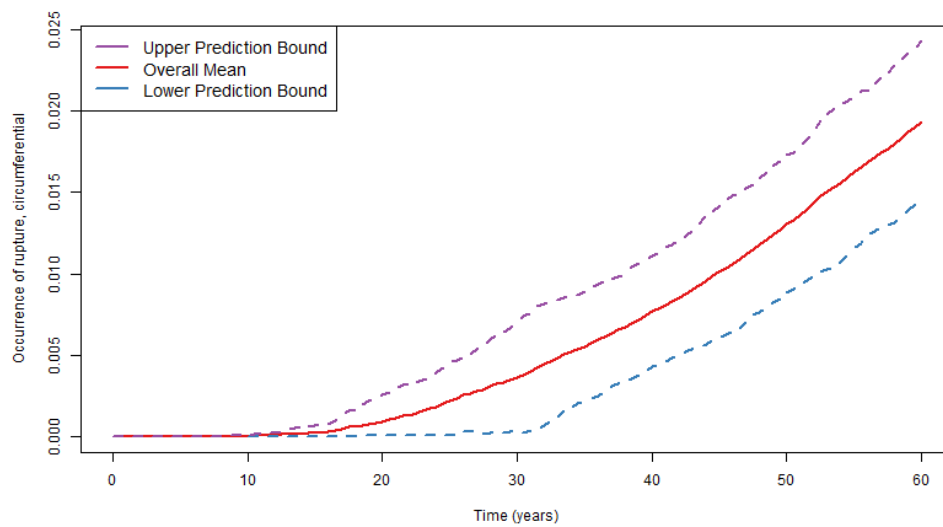


Figure 66: 95% prediction interval over mean probability of circumferential rupture for five replicates for Scenario 3.

3.4 Scenario 4

3.4.1 Scenario Summary

The results and conclusions of a detailed analysis of sampling schemes and their impact on an understanding of the uncertainty found in the simulation results of Scenario 4 are given in this section.

Scenario 4 uses the following defining options in the xLPR code:

- PWSCC flaw initiation
- Circumferential and axial flaw orientation
- PWSCC flaw growth
- Stress mitigation using mechanical stress improvement process (MSIP) at approximately 20 years

The outputs of interest defined for this scenario for both circumferential and axial cracks are:

- Probability of occurrence of crack
- Probability of occurrence of leak
- Probability of occurrence of rupture
- Total leak rate
- Maximum crack depth
- Maximum crack inner half-length
- Maximum crack outer half-length

The comprehensive list of simulations that were used in the analysis of this scenario is shown below. A constant aleatory sample size of 50 was used for each of the sampling schemes considered.

1. 100/50 (epistemic/aleatory sample size) with simple random sampling (SRS)
2. 500/50 with SRS
3. 1000/50 with Latin hypercube sampling (LHS)
4. 1000/50 with LHS and importance sampling (IS) on p2543 (multiplier on DM1 proportionality constant A) with target quantile of 0.95
5. 1000/50 with SRS and IS on p2543 with target quantile of 0.95

The results of the analysis of these sampling schemes are summarized as follows. Similar to Scenario 3, the regression analyses indicate that the same variable, the multiplier on Direct Model 1 (DM1) proportionality constant A for the PWSCC crack initiation module, hereafter referred to as p2543, was the most important input for explaining variability in all of the outputs. Because of this continuity and for the sake of brevity, the only result shown in Section 3.4.2 is the probability of occurrence of circumferential rupture. Overall, results confirm that increasing the epistemic sample size decreases the uncertainty in the mean of the output (e.g., probability of occurrence of circumferential rupture) as expected.

Compared to Scenario 3, stress mitigation at 20 years has a significant effect on the results. Behaviors of axial and circumferential cracks vary slightly after mitigation is implemented. For axial cracks, the probabilities and crack response results remain constant following the prescribed

mitigation. These results are attributed to the reduced post-mitigation compressive hoop stress WRS following mitigation. For circumferential cracks, the probabilities and crack response results continue to increase after mitigation, but the response is significantly dampened. Similar to the post-mitigation hoop stress, the post-mitigation axial WRS is reduced, but not as significantly as the hoop stress. Although the stress has been reduced, it is still at a value that would suggest continued crack initiation but at a much slower rate. Based on these results, it can be concluded that MSIP mitigation is an effective measure for stopping axial crack formation and growth and significantly affecting the rate of circumferential crack formation and growth.

3.4.2 Analysis Progression

Scenario 3 (Section 3.3.2) provided a comprehensive progression of the different sampling options available and a thorough analysis from run-to-run. The Scenario 4 analysis progression is not as comprehensive as Scenario 3, but highlights key aspects of the different sampling options to imitate the analysis progression provided in Section 3.3.2. For a thorough analysis of sampling options, the reader is referred back to Section 3.3.2. This section for Scenario 4 instead shows the run-to-run progression to demonstrate convergence of the solutions.

Runs 1 and 2 were used to determine if epistemic sample sizes have an impact on result uncertainty. Figure 67 plots the mean probability of occurrence of circumferential rupture for Run 1 and Run 2, along with the 95% confidence intervals (CIs), calculated using the bootstrap method (Section 2.2.2.2), which describes the sampling uncertainty associated with the mean. Following MSIP stress mitigation at 20 years, the probability continues to increase and approaches a steady value of ~ 0.0035 . The occurrence of rupture continues to increase because circumferential cracks continue to grow, albeit slower due to MSIP mitigation. The effect of MSIP on results is discussed further in Section 3.4.6. As seen in Figure 67, for a sample scheme of 500/50 SRS the mean is smoother and the width of the CIs is narrower. The CI width is better visualized in Figure 68. From Run 1 to Run 2, the width of the CIs decreases by ~ 0.02 .

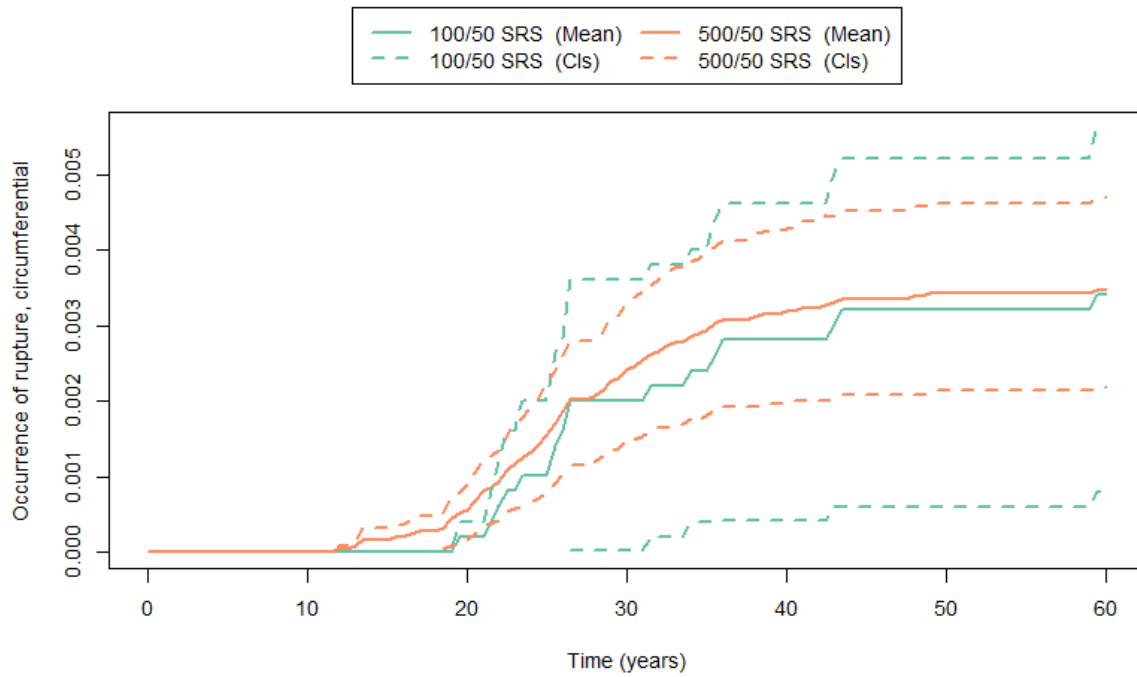


Figure 67: Mean probability of occurrence of circumferential rupture (solid line) and 95% confidence intervals (dashed lines) for Scenario 4, Run 1 (green) and Run 2 (orange).

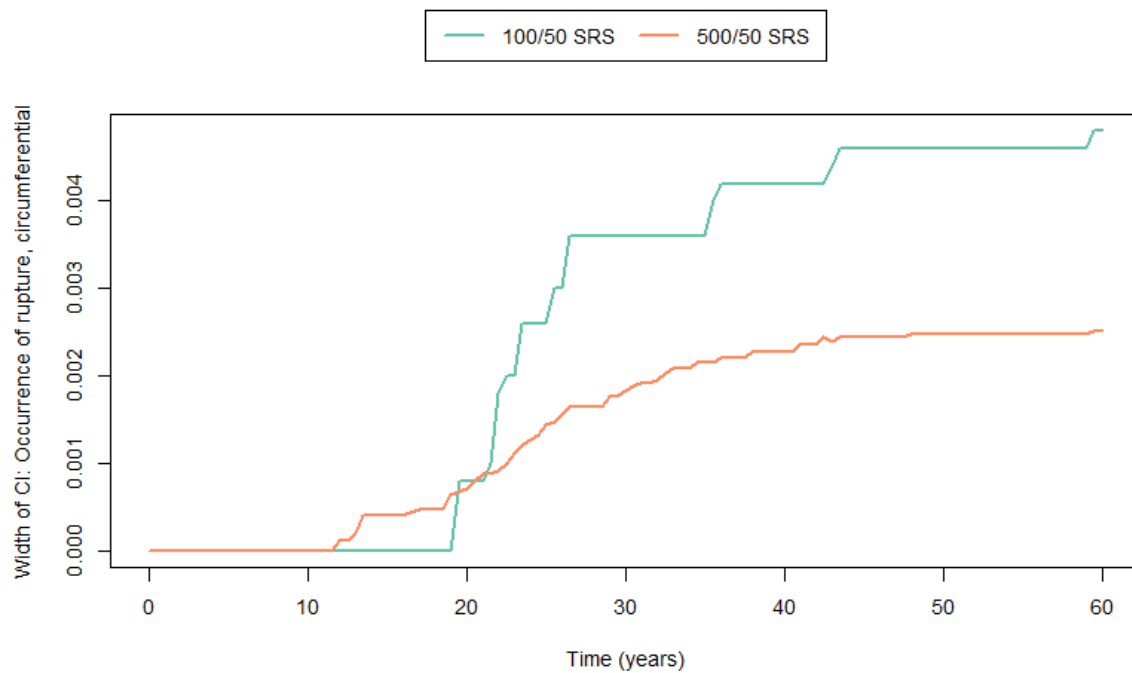


Figure 68: Width of 95% confidence interval for probability of occurrence of circumferential rupture for Scenario 4, Run 1 (green) and Run 2 (orange).

Based on the results seen in Figure 67 and Figure 68, and the analysis progression described in Section 3.3.2, the next progression in the analysis was to increase the epistemic sample size to 1000 and sample using LHS. The justification to move from SRS to LHS is described in Section 4.3. The results for Run 3 are plotted in Figure 69. The LHS confidence bounds presented in Figure 69 are calculated as SRS confidence bounds as described in the methods section and thus cannot serve as direct evidence comparing these two sampling methods (i.e., LHS and SRS).

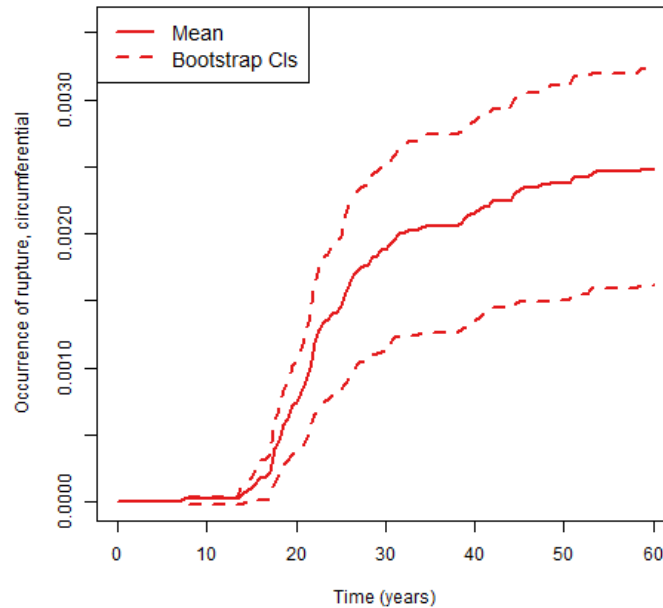


Figure 69: Mean probability of occurrence of circumferential rupture (solid line) and 95% confidence intervals (dashed lines) for Scenario 4, Run 3.

The results of preliminary runs (Runs 1 through 3) identify variable p2543, the multiplier on DM1 proportionality constant A, as the most important variable to all the scenario outputs. This importance can be seen both by a visual inspection of the relationship between this variable and the output of interest and through the use of a rank regression analysis.

Figure 70 shows the scatter plots for this variable and the probability of occurrence of circumferential rupture for Runs 1 through 3.

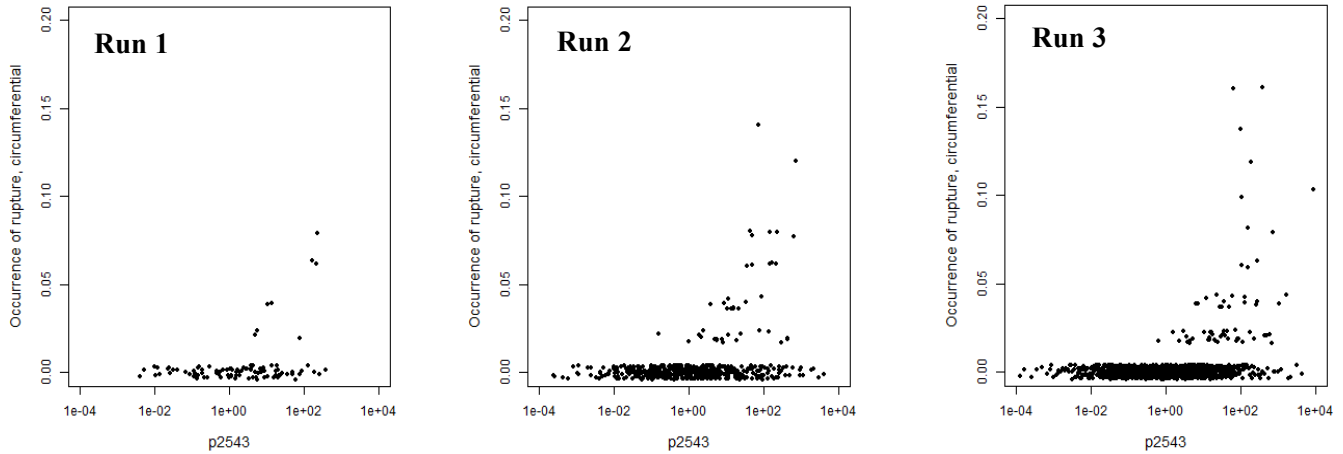


Figure 70: Scatter plots for variable p2543 (multiplier on DM1 proportionality constant A) and the probability of occurrence for circumferential rupture for Scenario 4, Run 1 (left), Run 2 (center), and Run 3 (right).

The scatterplots for these variables show that p2543 has a definitive relationship with the output of interest as higher values of this variable correlate to higher occurrences of circumferential rupture. A rank regression analysis was used to confirm the importance of variable p2543 to all the outputs of interest. The results of this rank regression analysis shown in Table 5 compare the standardized rank regression coefficient (SRRC) for the five most significant variables for each result found in Run 3. The probability of occurrence of axial rupture is not shown in this table because there were not any axial ruptures observed for this scenario.

The SRRC results for probability of circumferential leak and probability of circumferential rupture are identical because the probability at year 60 for these two outputs are also identical and only the value at year 60 is used to determine the SRRC.

Table 5: Summary of the stepwise rank regression analysis results for all outputs at year 60 for Scenario 4, Run 3

Variable Identifier	Variable Name	SRRC	Variable Identifier	Variable Name	SRRC
Probability of Occurrence of Axial Cracks – $R^2 = 0.807$			Maximum Axial Crack Depth – $R^2 = 0.771$		
p2543	Multiplier proport. Const. A (DM1)	0.885	p2543	Multiplier proport. Const. A (DM1)	0.843
p4350	Hoop WRS Pre-mitigation	-0.184	p4350	Hoop WRS Pre-mitigation	-0.234
p5106	Pre-Mitigation: depth-sizing slope term, b (circ)	0.092	p5106	Pre-Mitigation: depth-sizing slope term, b (circ)	0.094
p5105	Pre-Mitigation: depth-sizing bias term, a (circ)	0.060	p1102	Pipe Wall Thickness	-0.068
p9002	Surface crack dist. rule modifier	-0.042	p5105	Pre-Mitigation: depth-sizing bias term, a (circ)	0.065
Probability of Occurrence of Circumferential Cracks – $R^2 = 0.769$			Maximum Circumferential Crack Depth – $R^2 = 0.751$		
p2543	Multiplier proport. Const. A (DM1)	0.873	p2543	Multiplier proport. Const. A (DM1)	0.852
p4352	Axial WRS Pre-mitigation	-0.077	p4352	Axial WRS Pre-mitigation	-0.100
p2306	Material Init J-Resistance, Jic	0.035	p5401	Intercept, B0 (circ)	0.068
p4353	Axial WRS Post-mitigation	-0.033	p1102	Pipe Wall Thickness	-0.065
p1102	Pipe Wall Thickness	-0.031	p5402	Slope, B1 (circ)	0.053
Probability of Axial Leak – $R^2 = 0.251$			Maximum Axial Inner Half-Length – $R^2 = 0.796$		
p2543	Multiplier proport. Const. A (DM1)	0.364	p2543	Multiplier proport. Const. A (DM1)	0.871
p2592	Comp-to-Comp Variab Factor, fcomp	0.231	p4350	Hoop WRS Pre-mitigation	-0.220
p2594	Peak-to-Valley ECP Ratio - 1, P-1	-0.187	p5106	b (circ)	0.095
p2595	Charact Width of Peak vs ECP, c	0.170	p5105	a (circ)	0.066
p5401	Intercept, B0 (circ)	0.115	p2591	Activation Energy, Qg	0.048
Probability of Circumferential Leak – $R^2 = 0.221$			Maximum Circumferential Inner Half-Length – $R^2 = 0.745$		
p2543	Multiplier proport. Const. A (DM1)	0.322	p2543	Multiplier proport. Const. A (DM1)	0.845
p1102	Pipe Wall Thickness	-0.188	p4352	Axial WRS Pre-mitigation	-0.106
p2592	Comp-to-Comp Variab Factor, fcomp	0.167	p5401	Intercept, B0 (circ)	0.071
p4352	Axial WRS Pre-mitigation	-0.129	p1102	Pipe Wall Thickness	-0.055
p2594	Peak-to-Valley ECP Ratio - 1, P-1	-0.114	p2592	Comp-to-Comp Variab Factor, fcomp	0.053
Probability of Circumferential Rupture – $R^2 = 0.221$			Maximum Axial Outer Half-Length – $R^2 = 0.697$		
p2543	Multiplier proport. Const. A (DM1)	0.322	p2543	Multiplier proport. Const. A (DM1)	0.364
p1102	Pipe Wall Thickness	-0.188	p2592	Comp-to-Comp Variab Factor, fcomp	0.231
p2592	Comp-to-Comp Variab Factor, fcomp	0.167	p2594	Peak-to-Valley ECP Ratio - 1, P-1	-0.187
p4352	Axial WRS Pre-mitigation	-0.129	p2595	Charact Width of Peak vs ECP, c	0.170
p2594	Peak-to-Valley ECP Ratio - 1, P-1	-0.114	p5401	Intercept, B0 (circ)	0.115
Total Leak Rate – $R^2 = 0.313$			Maximum Circumferential Outer Half-Length – $R^2 = 0.221$		
p2543	Multiplier proport. Const. A (DM1)	0.414	p2543	Multiplier proport. Const. A (DM1)	0.323
p2592	Comp-to-Comp Variab Factor, fcomp	0.250	p1102	Pipe Wall Thickness	-0.187
p2594	Peak-to-Valley ECP Ratio - 1, P-1	-0.160	p2592	Comp-to-Comp Variab Factor, fcomp	0.167
p2595	Charact Width of Peak vs ECP, c	0.156	p4352	Axial WRS Pre-mitigation	-0.131
p1102	Pipe Wall Thickness	-0.133	p2594	Peak-to-Valley ECP Ratio - 1, P-1	-0.115

Table 5 indicates that variable p2543 is the most significant variable because the magnitude of its SRRC is much higher than all of the other sampled variables used for this scenario across all the outputs that are being analyzed. This variable is used in the PWSCC crack initiation model (DM1) in xLPR. After p2543, there are no consistently recurring variables that can be identified as the second most important variable. However, variables related to PWSCC crack growth properties (such as p2592, Comp-to-Comp Variab Factor, fcomp) and WRS properties for both pre- and post-mitigation are consistently identified as being one of the five most important variables.

Although p2543 was found to be the most important variable related to the outputs of interest, the upper tail of the input distribution for p2543, which is the region most likely to result in cracks, is not well sampled using traditional SRS or LHS. Because this variable was found to be strongly correlated with the results of interest, importance sampling was applied to p2543 at the 95th quantile to improve the coverage at the upper tail of the distribution. This sampling selection led to a larger number of circumferential ruptures, as shown in Figure 71 which compares Runs 3 and 4. This

result confirms both the importance of variable p2543 to the results of this scenario and that the importance sampling was able to generate a greater numbers of events for a fixed number of samples.

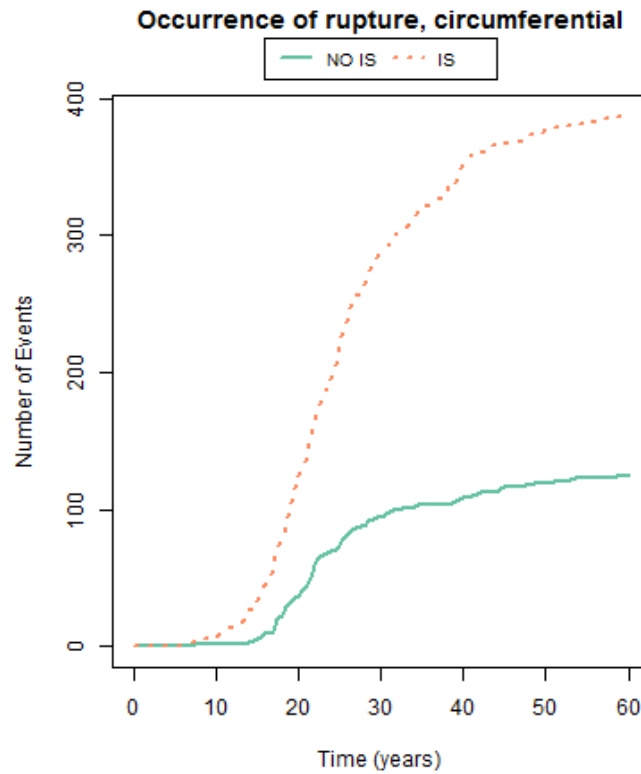


Figure 71: Number of circumferential rupture events occurring for Scenario 4, Run 3 (NO IS) and Run 4 (IS).

Figure 72 and Figure 73 show the effects of importance sampling for 1000/50 sampling size using LHS. The width of the confidence interval for occurrence of circumferential rupture does not decrease at 60 years using importance sampling. A separate study on the importance of the quantile selected for importance sampling is provided in Section 4.2.

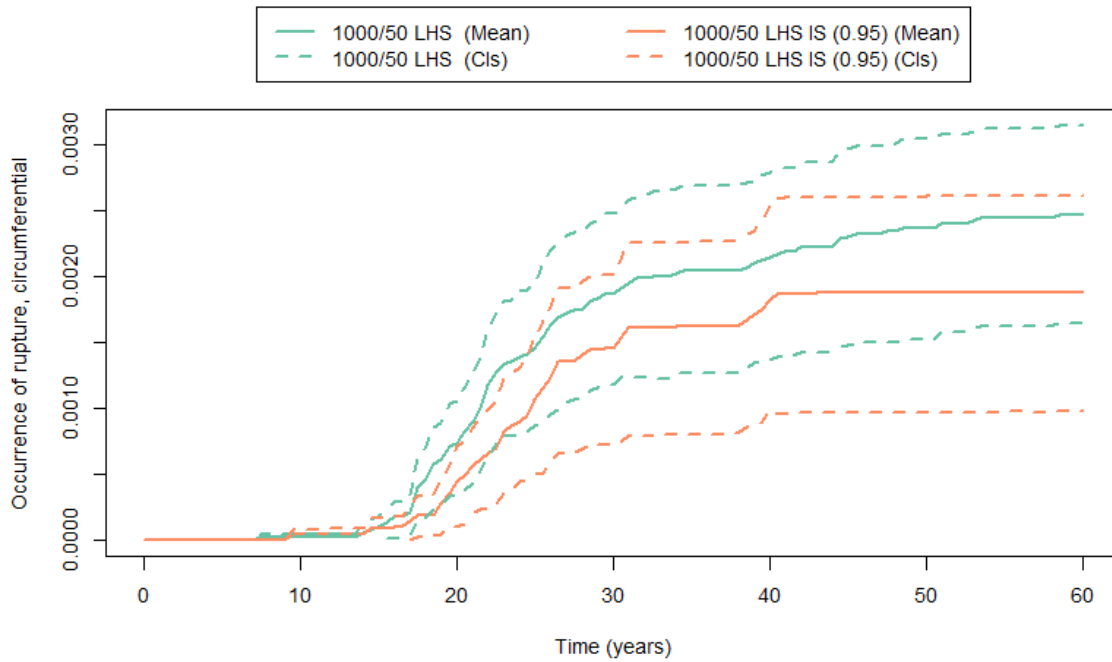


Figure 72: Mean probability of occurrence of circumferential rupture (solid line) and 95% confidence intervals (dashed lines) for Scenario 4, Run 3 (green) and Run 4 (orange).

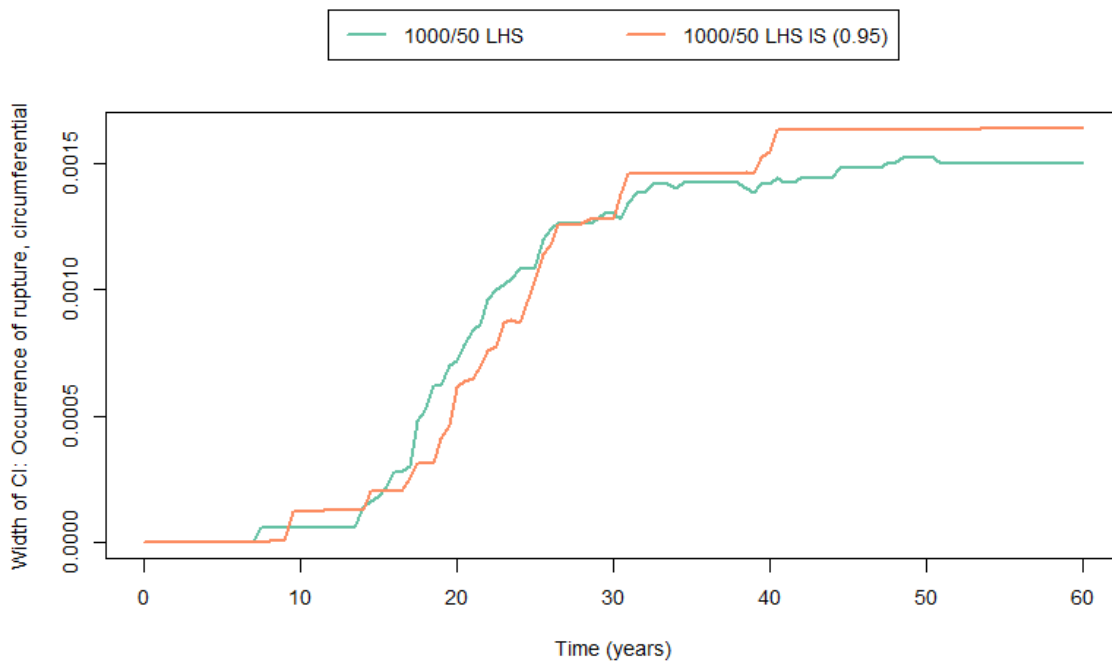


Figure 73: Width of 95% confidence interval for probability of occurrence of circumferential rupture for Scenario 4, Run 3 (green) and Run 4 (orange).

3.4.3 Overall Results

The impact of sampling schemes will vary based on the output of interest. Figure 74 shows the impact of sampling schemes on the mean probability of occurrence of circumferential rupture for Runs 1 through 5. While Figure 74 shows that sampling schemes using importance sampling and/or LHS are approximately equal with respect to the magnitude of the mean over the simulation time, Figure 75 better illustrates the convergence of these sampling schemes for the mean probability of occurrence of circumferential cracks. The divergent results found using SRS alone can be attributed to the dependence of this sampling scheme on the selected random seed, as is studied in Section 4.4.

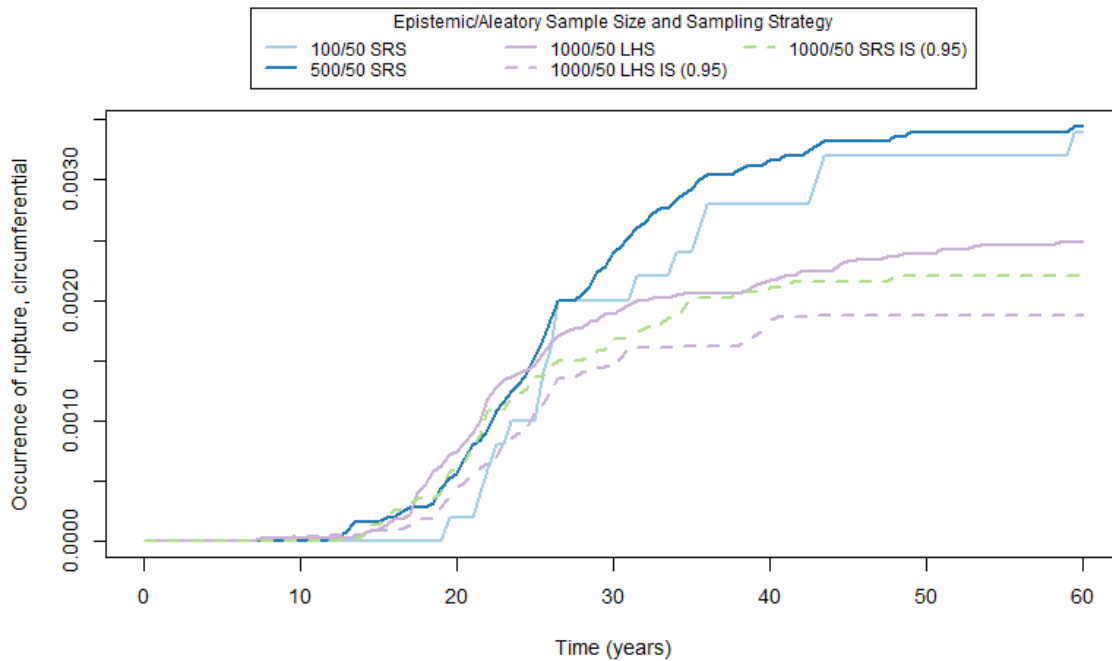


Figure 74: Mean probability of occurrence of circumferential rupture for Scenario 4, Runs 1 through 5.

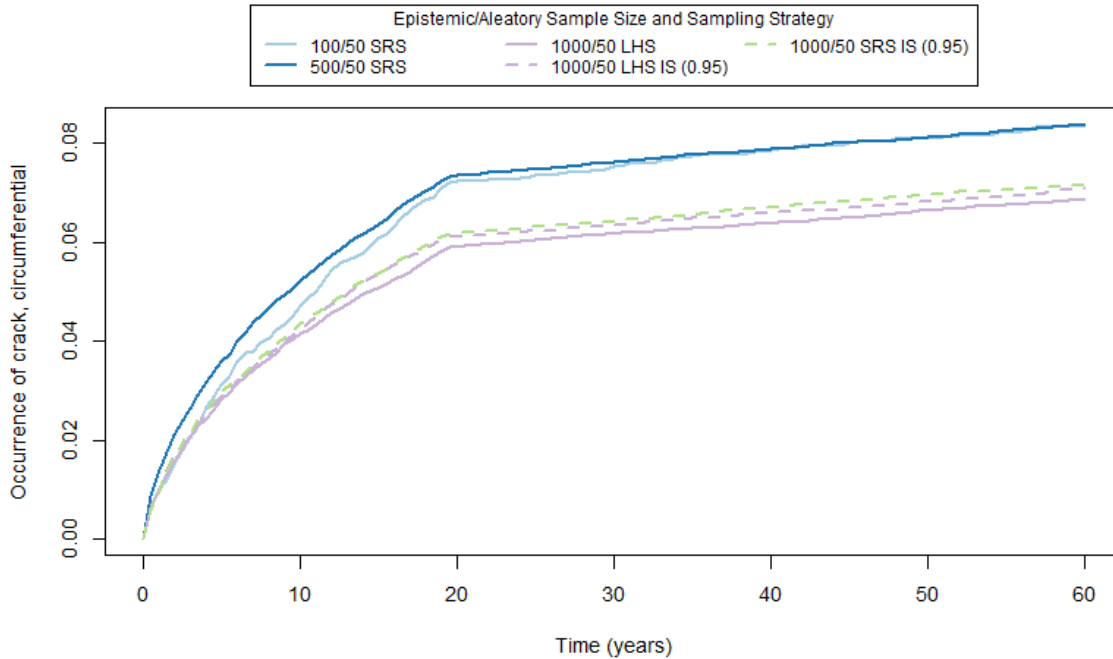


Figure 75: Mean probability of occurrence of circumferential cracks for Scenario 4, Runs 1 through 5.

3.4.4 Epistemic Uncertainty Analysis

An uncertainty analysis was conducted for the sampling scheme used in Run 4 (1000 epistemic and 50 aleatory samples and LHS with IS on p2543) with the goal of characterizing epistemic uncertainty in the probabilities of occurrence.

Plots of the estimated probabilities of occurrence of circumferential crack and rupture are displayed in Figure 76 and Figure 77, respectively. These figures show the probability of occurrence for each epistemic realization (grey lines), the mean (red line), and the 5th, 50th, and 95th percentiles (blue, green, and purple lines, respectively). As noted in Section 2.5.4 the percentiles shown here describe the variation found in the particular set of epistemic results.

Figure 76 shows that half of the epistemic realizations (i.e., the 50th percentile) have approximately a 2% chance of circumferential crack occurrence at year 60. Figure 77 shows that 5% (95th percentile) of epistemic realizations have an occurrence of circumferential rupture that is greater than 2%. The difference in the lower and upper quantiles for probability of circumferential rupture is small and up until year 39, as the 5th, 50th, and 95th percentiles are indistinguishable from one another.

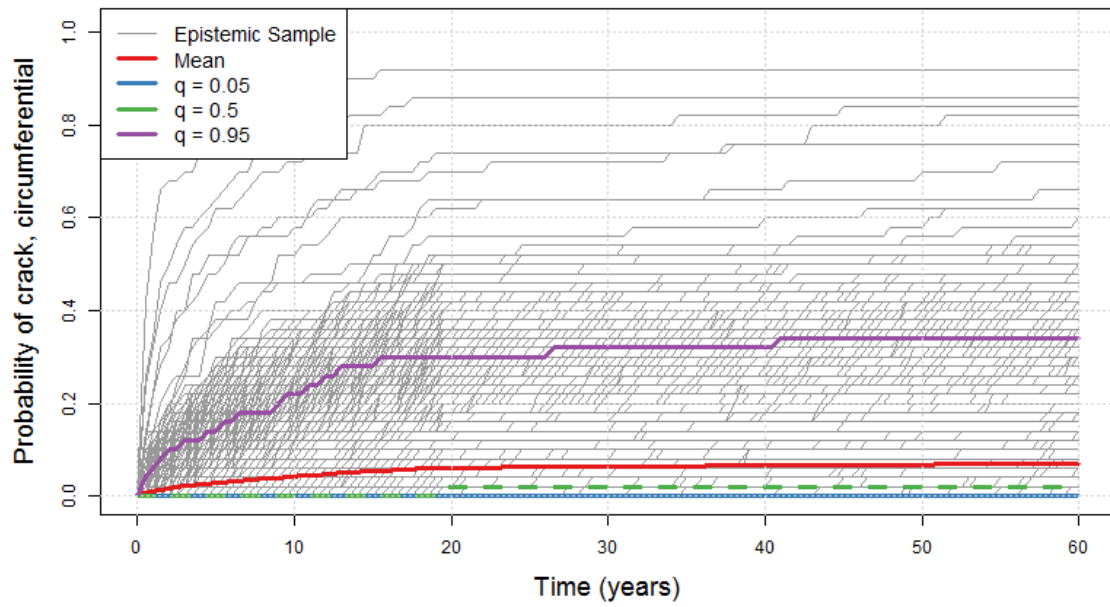


Figure 76: Probability of occurrence of circumferential crack for each epistemic realization (grey), the mean (red), and the 5th (blue), 50th (green), and 95th (purple) percentiles for Scenario 4, Run 4.

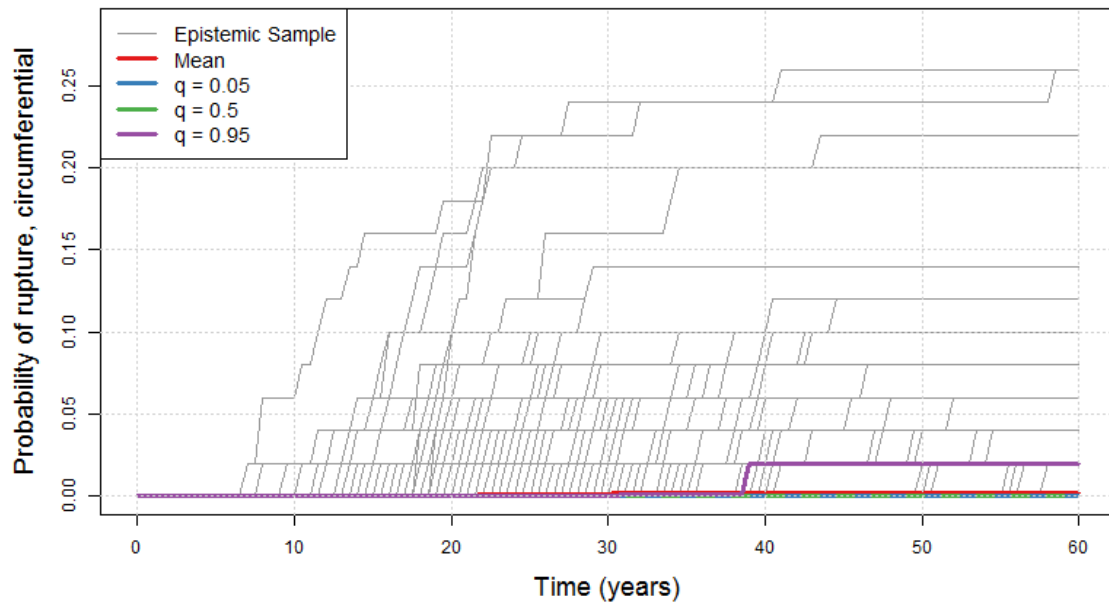


Figure 77: Probability of occurrence of circumferential rupture for each epistemic realization (grey), the mean (red), and the 5th (blue), 50th (green), and 95th (purple) percentiles for Scenario 4, Run 4.

Figure 78 compares the mean and 95th percentile of the probability of occurrence for circumferential leaks and ruptures. As expected, circumferential leaks precede ruptures but comparing the mean probabilities, circumferential leaks are quickly followed by ruptures. The delay between occurrence of leaks and ruptures are greater for the 95th percentile but by year 40, both events have equal probabilities of occurrence.

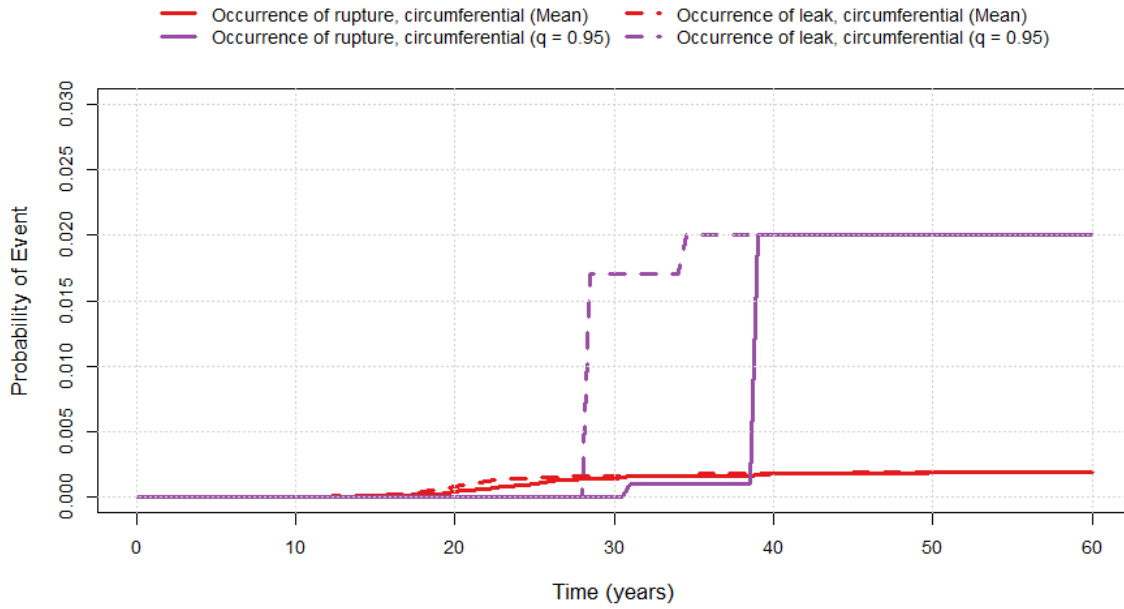


Figure 78: Mean (red) and 95th percentile (purple) for occurrence of circumferential leak (dashed) and rupture (solid) for Scenario 4, Run 4.

3.4.5 Convergence Analysis

The convergence of the sampling scheme for Scenario 4, Run 4 (1000 epistemic samples with LHS and IS) was analyzed using five replicate simulations with different random number seeds for both the epistemic and aleatory loops. Figure 79 shows the mean probability of occurrence of circumferential cracks for each replicate and Figure 80 shows the mean probability of occurrence of circumferential rupture for each replicate. A 95% prediction interval was created around the overall mean of the five replicates, as seen in Figure 81 and Figure 82 for occurrence of circumferential crack and rupture, respectively. The width of the 95% prediction interval is relatively narrow over the course of the simulation time for occurrence of circumferential cracks. However, the width of the 95% prediction interval for circumferential rupture is large; increasing the epistemic sample size would decrease variability in the mean estimate. Alternatively, in practice, if multiple convergence runs are available, we could average results across these 5 different runs to obtain a more precise estimate of the mean.

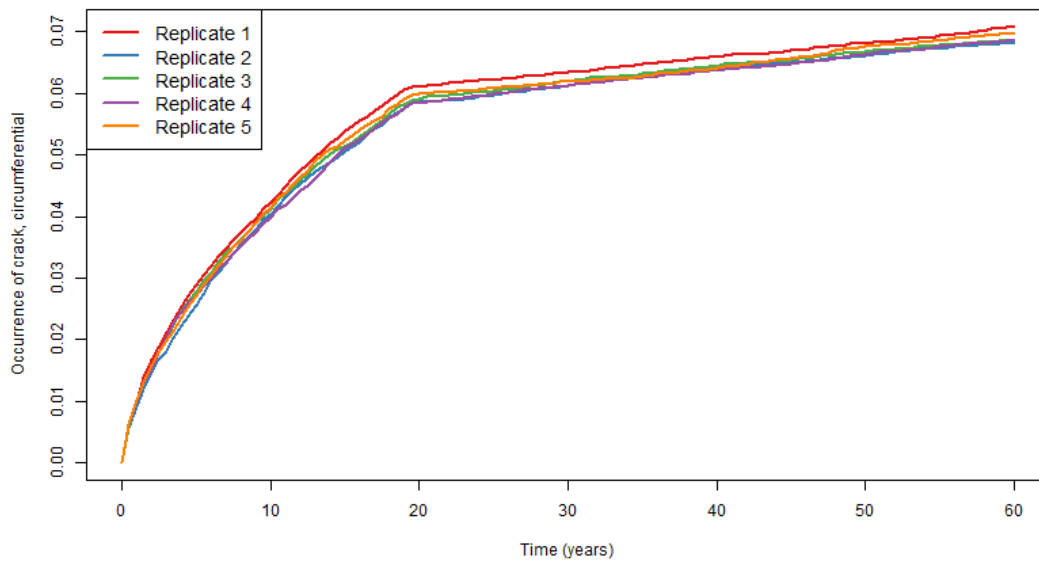


Figure 79: Comparison of estimated mean probability of occurrence of circumferential cracks for five convergence replicates for Scenario 4.

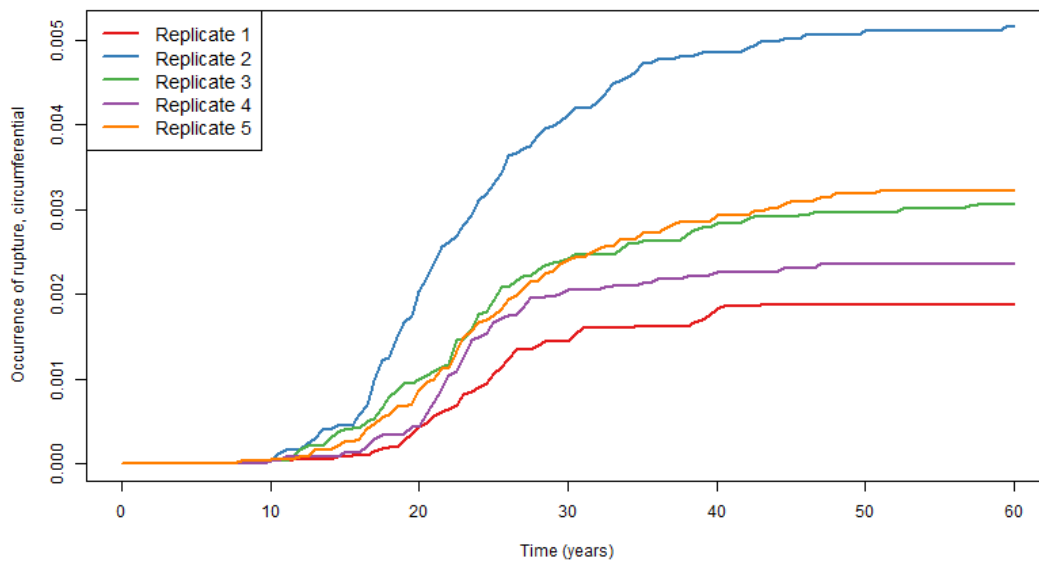


Figure 80: Comparison of estimated mean probability of occurrence of circumferential rupture for five convergence replicates for Scenario 4.

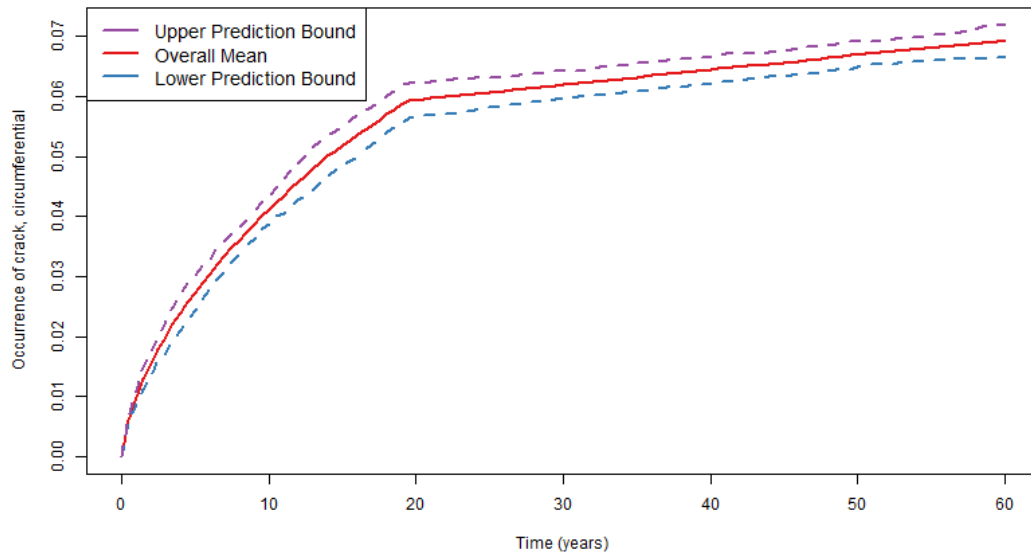


Figure 81: 95% prediction interval over mean probability of circumferential crack for five replicates for Scenario 4.

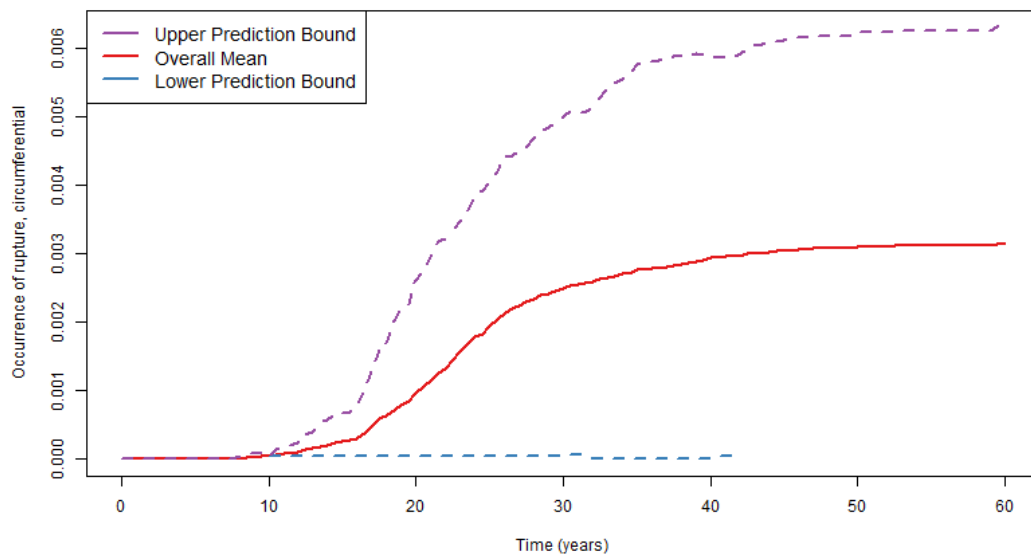


Figure 82: 95% prediction interval over mean probability of circumferential rupture for five replicates for Scenario 4.

3.4.6 Comparison to Scenario 3 Results

Scenario 4 runs were compared to Scenario 3 runs with corresponding sampling selections. To properly compare the two scenarios, Run 9 from Scenario 3 and Run 4 from Scenario 4 were used for the comparison as both runs represent sampling schemes of 1000 epistemic realizations with LHS and IS on p2543.

Scenario 3 and Scenario 4 comparisons are shown in Figure 83 through Figure 89 for selected results. After MSIP mitigation is implemented at year 20, the occurrence of axial cracks remains constant and the occurrence of circumferential cracks continues to increase but at a significantly slower rate. The reason for occurrence of axial cracks remaining constant is that the resulting post-mitigation compressive hoop WRS is significantly lowered such that no more axial cracks initiate. The post-mitigation compressive axial WRS is also lowered following MSIP mitigation, but not enough to completely halt circumferential crack formation.

Occurrence of axial and circumferential leak results are similar to crack results as seen in Figure 84. Figure 85 provides a comparison for occurrence of circumferential rupture (note that axial rupture is not predicted for Scenario 3 or Scenario 4). Crack property results for maximum crack half-length (iHL), maximum outer half-length (oHL), and maximum crack depth are shown in Figure 86, Figure 87, and Figure 88, respectively. MSIP mitigation is particularly effective at limiting crack growth as the crack properties see marginal increases from year 20 to year 60. Additionally, the total leak rate for Scenario 3 and Scenario 4 are compared in Figure 89. When MSIP mitigation is applied at 20 years (Scenario 4), the total leak rate at 60 years is about an order of magnitude lower compared to no mitigation (Scenario 3). Note that the total leak rate is not affected by repairs or leak detection.

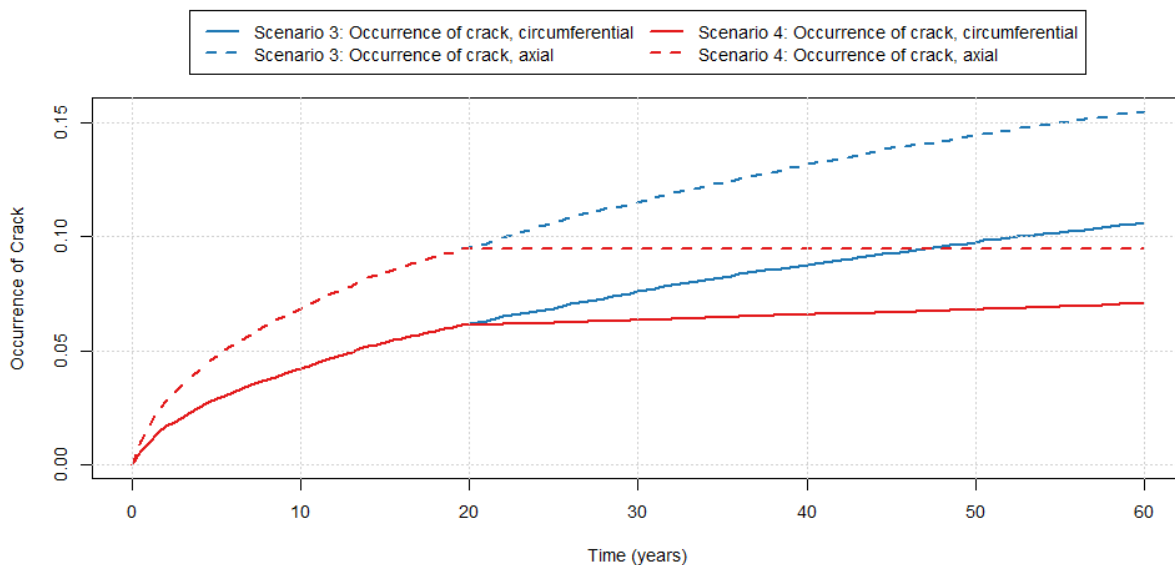


Figure 83: Mean probability of axial (dashed lines) and circumferential (solid lines) crack occurrence for Scenario 3 (blue lines) and Scenario 4 (red lines).

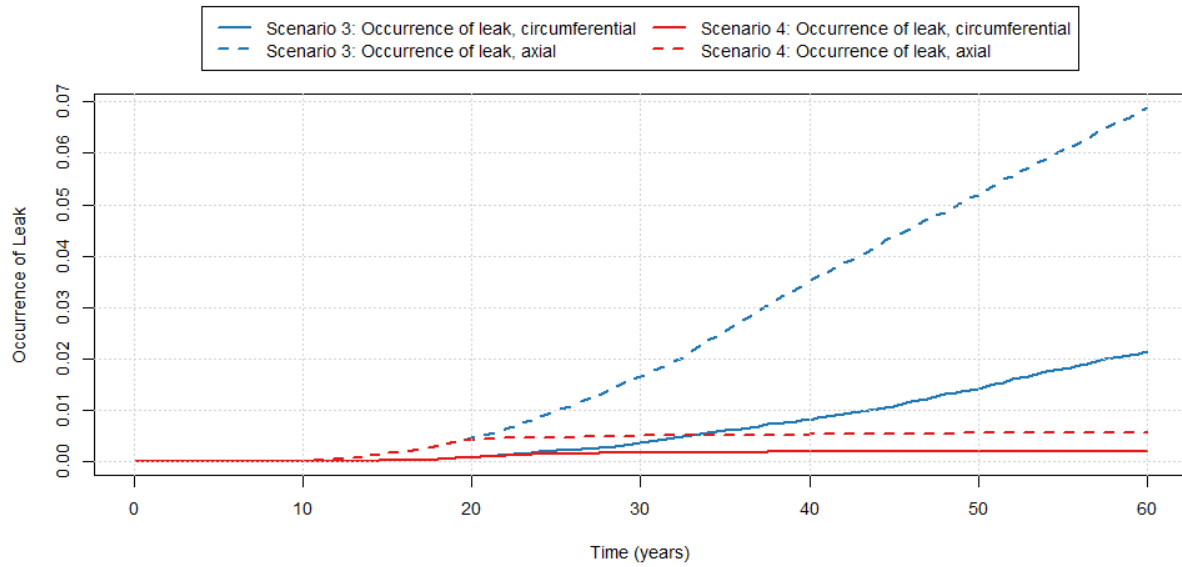


Figure 84: Mean probability of axial (dashed lines) and circumferential (solid lines) leak occurrence for Scenario 3 (blue lines) and Scenario 4 (red lines).

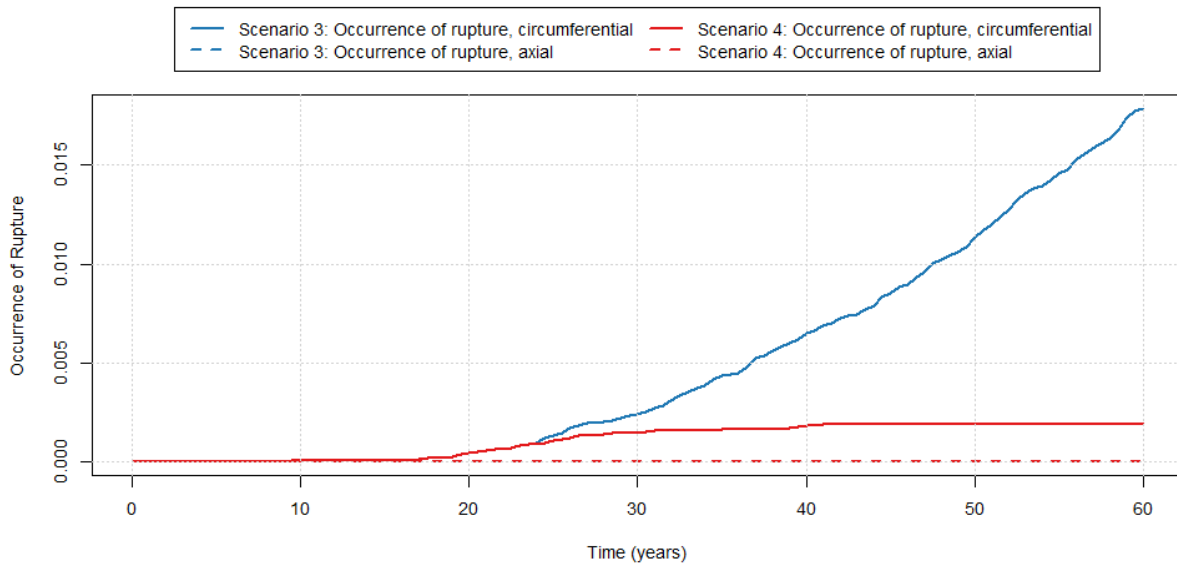


Figure 85: Mean probability of axial (dashed lines) and circumferential (solid lines) rupture occurrence for Scenario 3 (blue lines) and Scenario 4 (red lines).

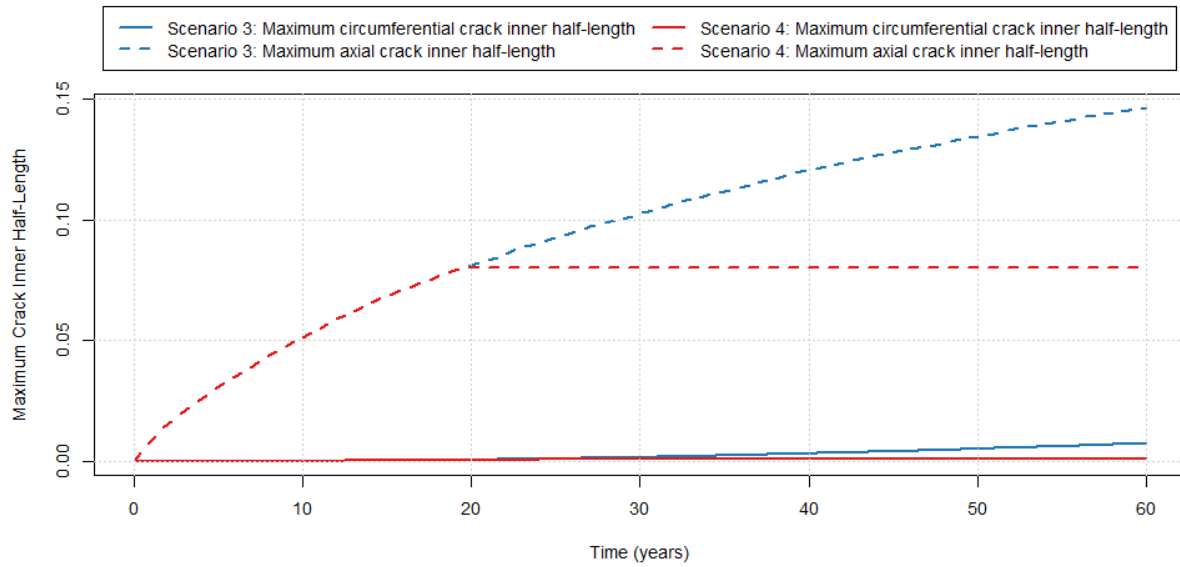


Figure 86: Mean maximum axial (dashed lines) and circumferential (solid lines) iHL for Scenario 3 (blue lines) and Scenario 4 (red lines).

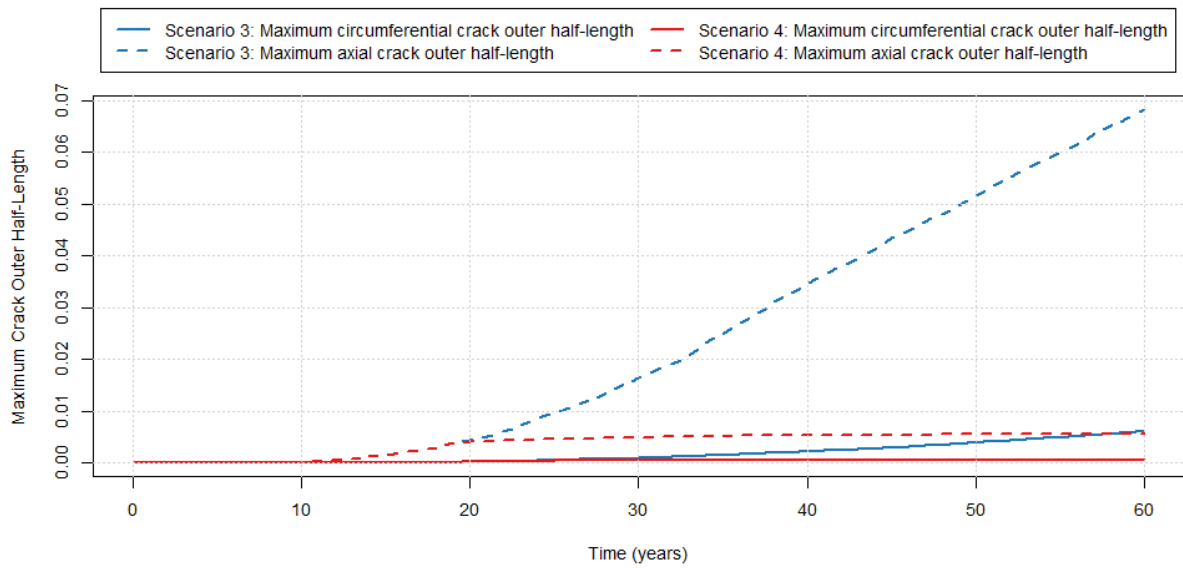


Figure 87: Mean maximum axial (dashed lines) and circumferential (solid lines) oHL for Scenario 3 (blue lines) and Scenario 4 (red lines).

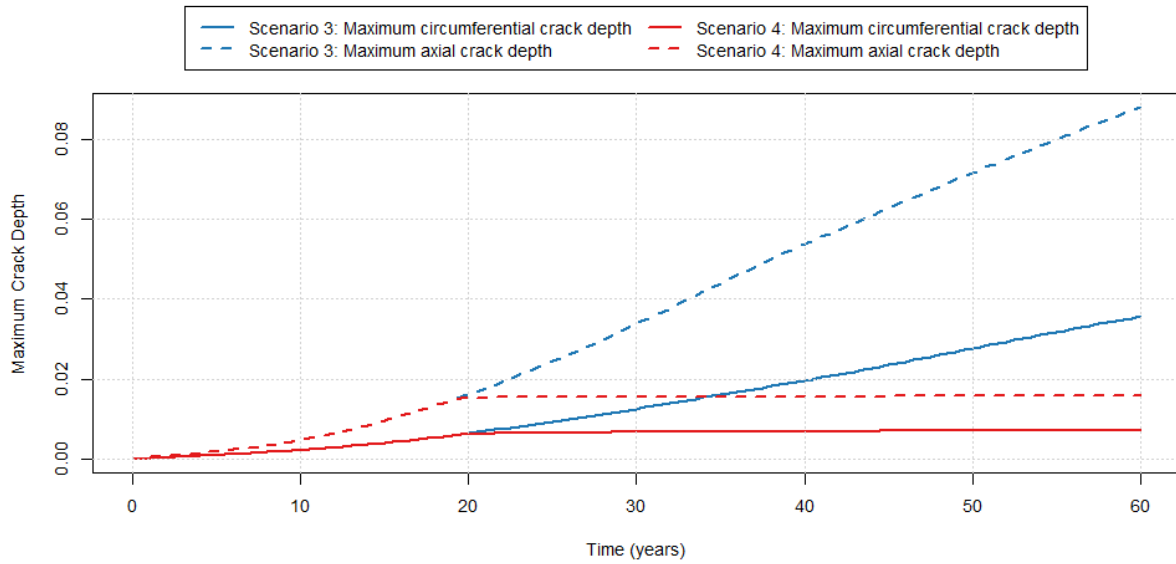


Figure 88: Mean maximum axial (dashed lines) and circumferential (solid lines) crack depth for Scenario 3 (blue lines) and Scenario 4 (red lines).

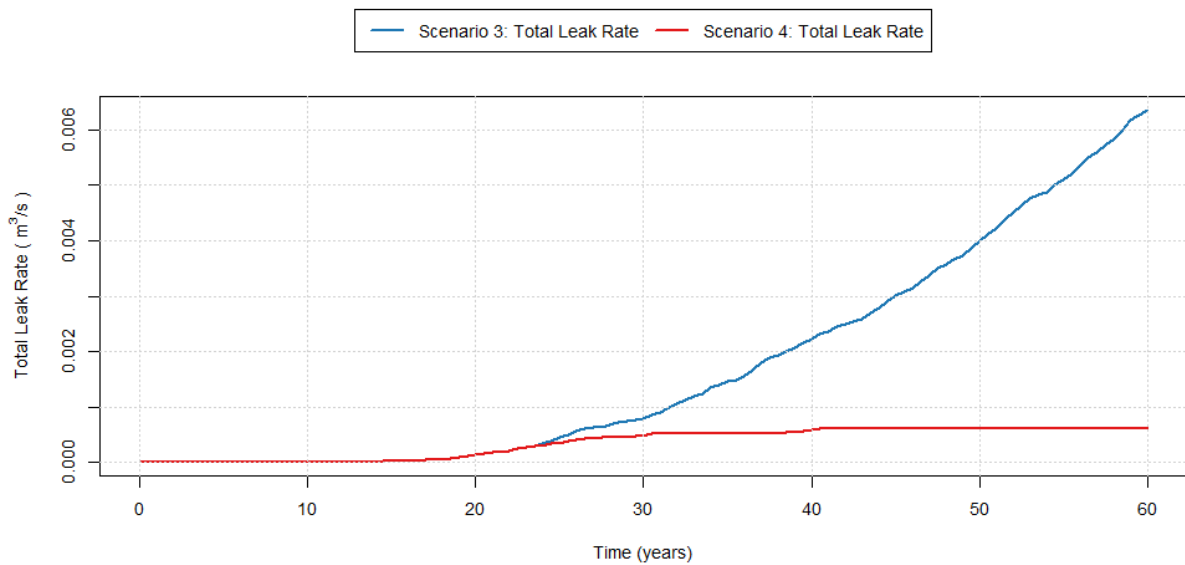


Figure 89: Mean total leak rate for Scenario 3 (blue line) and Scenario 4 (red line).

3.5 Scenario 5

3.5.1 Scenario Summary

The results and conclusions of a detailed analysis of sampling schemes and their impact on an understanding of the uncertainty found in the simulation results of Scenario 5 are given in this section.

Scenario 5 uses the following defining options in the xLPR code:

- PWSCC flaw initiation
- Circumferential and axial flaw orientation
- PWSCC flaw growth
- Stress mitigation using mechanical stress improvement process (MSIP) at approximately 40 years

The outputs of interest defined for this scenario for both circumferential and axial cracks are:

- Probability of occurrence of crack
- Probability of occurrence of leak
- Probability of occurrence of rupture
- Total leak rate
- Maximum crack depth
- Maximum crack inner half-length
- Maximum crack outer half-length

The comprehensive list of simulations that were used in the analysis of this scenario is shown below. A constant aleatory sample size of 50 was used for each of the sampling schemes considered.

1. 100/50 (epistemic/aleatory sample size) with simple random sampling (SRS)
2. 500/50 with SRS
3. 1000/50 with Latin hypercube sampling (LHS)
4. 1000/50 with LHS and importance sampling (IS) on p2543 (multiplier on DM1 proportionality constant A) with target quantile of 0.95
5. 1000/50 with SRS and IS on p2543 with target quantile of 0.95

The results of the analysis of these sampling schemes are summarized as follows. Like Scenarios 3 and 4, the results of the regression analyses indicate that the same variable, the multiplier on Direct Model 1 (DM1) proportionality constant A for the PWSCC crack initiation module, hereafter referred to as p2543, was the most important input for explaining the variability in all of the outputs. Because of this continuity and for the sake of brevity, the only output shown in Section 3.5.2 is the probability of occurrence of circumferential rupture. Overall, results confirm that increasing the epistemic sample size decreases the uncertainty in the mean of the output, as expected. If higher precision is required for low probability events, such as occurrence of circumferential rupture, then the number of epistemic realizations should be increased.

Stress mitigation using MSIP is an effective mitigation strategy for reducing the rate of crack formation. When MSIP mitigation is implemented, axial crack initiation is halted but axial cracks continue to grow, albeit at a significantly reduced rate. Circumferential crack formation and growth

continues following mitigation, but is significantly lower than the rates observed for Scenario 3. Implementing MSIP mitigation at 20 years versus 40 years reduces the probabilities and crack behavior substantially. The probabilities of occurrence of leak (both axial and circumferential), occurrence of circumferential rupture, and the total leak rate increased about six times more when MSIP mitigation was implemented at 40 years relative to mitigation at 20 years.

3.5.2 Analysis Progression

Scenario 3 (Section 3.3) provided a comprehensive progression of the different sampling options available and a thorough analysis from run-to-run. The Scenario 5 analysis progression is not as comprehensive as Scenario 3, but highlights key aspects of the different sampling options to imitate the analysis progression provided in Section 3.3. For a thorough analysis of sampling options, the reader is referred back to Section 3.3. This section for Scenario 5 instead shows the run-to-run progression to demonstrate convergence of the solutions.

Runs 1 and 2 were used to determine if epistemic sample sizes have an impact on result uncertainty. Figure 90 plots the mean probability of occurrence of axial cracks for Run 1 and Run 2, along with the 95% confidence intervals (CIs), calculated using the bootstrap method (Section 2.2.2.2), which describes the uncertainty associated with the mean. Following MSIP stress mitigation at 40 years, the rate at which the probability increases is slowed and approaches a plateaued value of ~ 0.016 . The effect of MSIP on results is discussed further in Section 3.5.6. As seen in Figure 90, the width of the CIs is narrower for the 500/50 SRS run, which is better visualized in Figure 91. From Run 1 to Run 2, the width of the CIs decreases by ~ 0.01 .

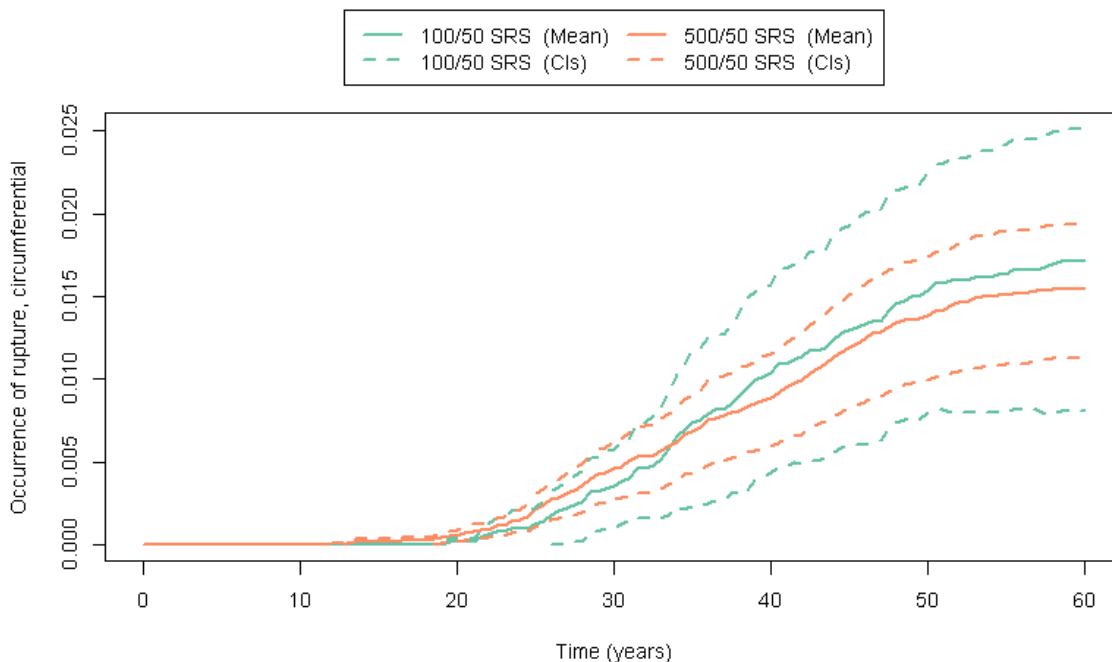


Figure 90: Mean probability of occurrence of circumferential rupture (solid line) and 95% confidence intervals (dashed lines) for Scenario 5, Run 1 (green) and Run 2 (orange).

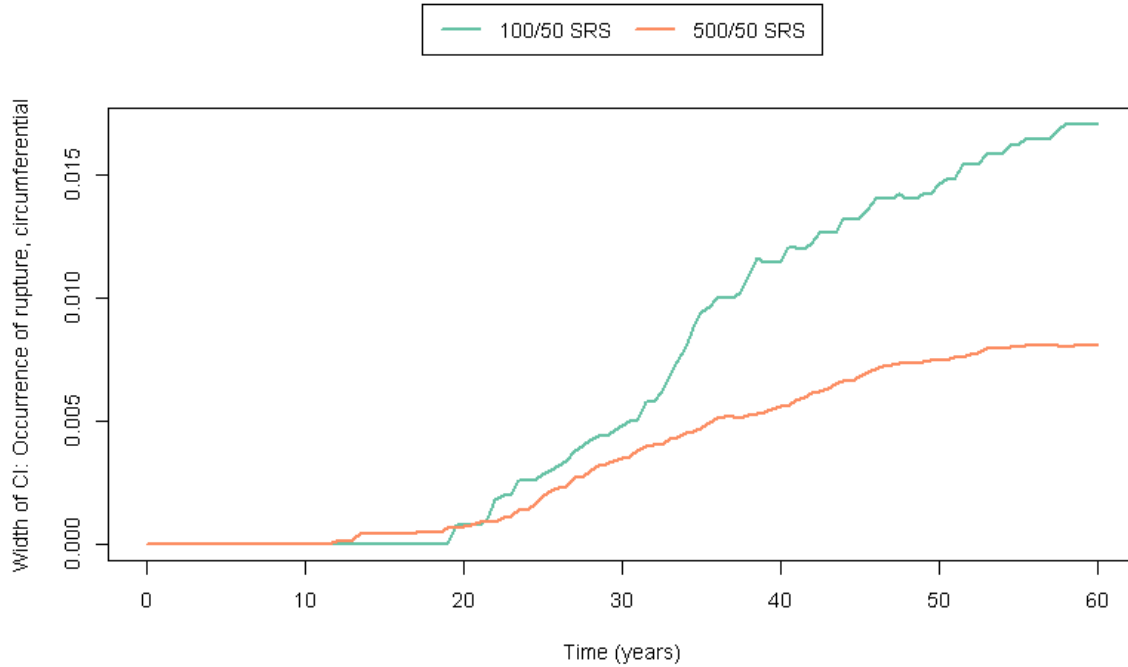


Figure 91: Width of 95% confidence interval for probability of occurrence of circumferential rupture for Scenario 5, Run 1 (green) and Run 2 (orange).

Based on the results seen in Figure 90 and Figure 91, and the analysis progression described in Section 3.3.2, the next progression in the analysis was to increase the epistemic sample size to 1000 and sample using LHS. The justification to move from SRS to LHS is described in Section 4.3. The results for Run 3 are plotted in Figure 92. The LHS confidence bounds presented in Figure 92 are calculated as SRS confidence bounds as described in the methods section and thus cannot serve as direct evidence comparing these two sampling methods (i.e., LHS and SRS).

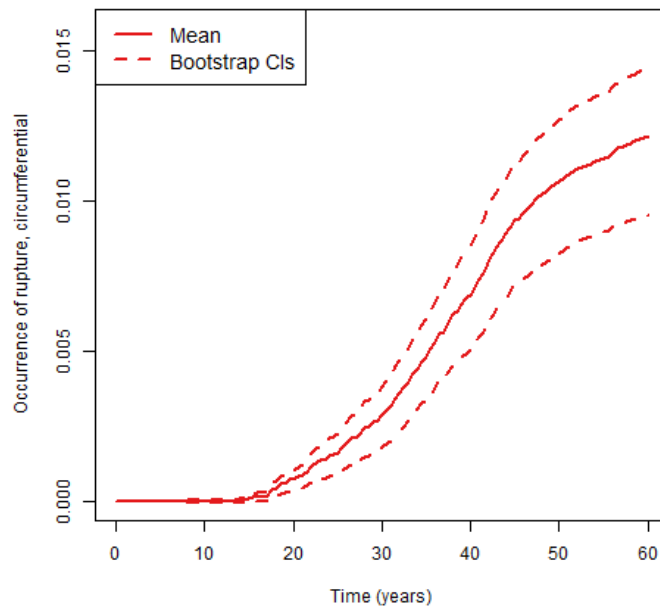


Figure 92: Mean probability of occurrence of circumferential rupture (solid line) and 95% confidence intervals (dashed lines) for Scenario 5, Run 3.

The results of preliminary runs (Runs 1 through 3) identify variable p2543, the multiplier on DM1 proportionality constant A, as the most important variable to all of the scenario outputs. This importance can be seen both by visual inspection of the relationship between this variable and the output of interest and through the use of rank regression analysis.

Figure 93 shows the scatter plots for this variable and the probability of occurrence of circumferential rupture for Runs 1 through 3.

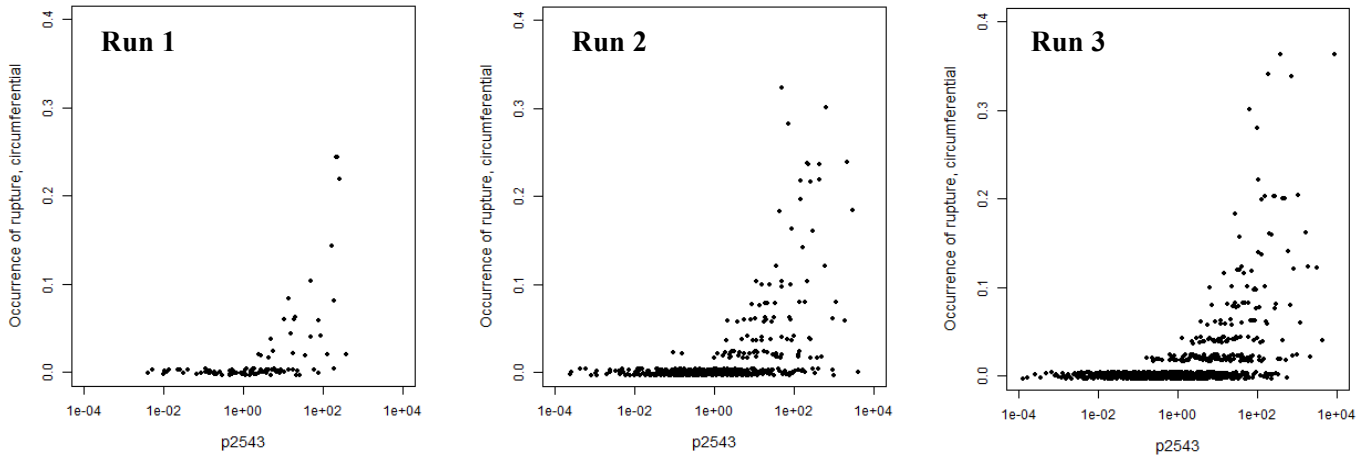


Figure 93: Scatter plots for variable p2543 (multiplier on DM1 proportionality constant A) and the probability of occurrence for circumferential rupture for Scenario 5, Run 1 (left), Run 2 (center), and Run 3 (right).

The scatterplots for these variables show that p2543 has a definitive relationship with the output of interest as higher values of this variable correlate to higher occurrences of circumferential rupture. A rank regression analysis was used to confirm the importance of variable p2543 to the outputs of interest. The results of this rank regression analysis shown in Table 6 compare the standardized rank regression coefficient (SRRC) for the five most significant variables for each result found in Run 3. The probability of occurrence of axial rupture is not shown in this table because there were not any axial ruptures observed for this scenario.

Table 6: Summary of the stepwise rank regression analysis results for all outputs at year 60 for Scenario 5, Run 3.

Variable Identifier	Variable Name	SRRC	Variable Identifier	Variable Name	SRRC
Probability of Occurrence of Axial Cracks – $R^2 = 0.858$			Maximum Axial Crack Depth – $R^2 = 0.797$		
p2543	Multiplier proport. Const. A (DM1)	0.909	p2543	Multiplier proport. Const. A (DM1)	0.844
p4350	Hoop WRS Pre-mitigation	-0.182	p4350	Hoop WRS Pre-mitigation	-0.234
p5401	Intercept, B0 (circ)	0.071	p2592	Comp-to-Comp Variab Factor, fcomp	0.087
p5402	Slope, B1 (circ)	0.064	p1102	Pipe Wall Thickness	-0.077
p5106	b (circ)	0.059	p5401	Intercept, B0 (circ)	0.075
Probability of Occurrence of Circumferential Cracks – $R^2 = 0.823$			Maximum Circumferential Crack Depth – $R^2 = 0.793$		
p2543	Multiplier proport. Const. A (DM1)	0.903	p2543	Multiplier proport. Const. A (DM1)	0.865
p4352	Axial WRS Pre-mitigation	-0.106	p4352	Axial WRS Pre-mitigation	-0.147
p1102	Pipe Wall Thickness	-0.029	p1102	Pipe Wall Thickness	-0.075
p9003	TW Crack Distance Rule Modifier	-0.026	p2592	Comp-to-Comp Variab Factor, fcomp	0.066
p5408	b (axial) Probability	0.021	p5105	a (circ)	0.055
Probability of Axial Leak – $R^2 = 0.540$			Maximum Axial Inner Half-Length – $R^2 = 0.847$		
p2543	Multiplier proport. Const. A (DM1)	0.629	p2543	Multiplier proport. Const. A (DM1)	0.894
p2592	Comp-to-Comp Variab Factor, fcomp	0.224	p4350	Hoop WRS Pre-mitigation	-0.218
p4350	Hoop WRS Pre-mitigation	-0.198	p5401	Intercept, B0 (circ)	0.075
p2595	Charact Width of Peak vs ECP, c	0.178	p5402	Slope, B1 (circ)	0.066
p2594	Peak-to-Valley ECP Ratio - 1, P-1	-0.155	p5106	b (circ)	0.065
Probability of Circumferential Leak – $R^2 = 0.422$			Maximum Circumferential Inner Half-Length – $R^2 = 0.780$		
p2543	Multiplier proport. Const. A (DM1)	0.536	p2543	Multiplier proport. Const. A (DM1)	0.857
p1102	Pipe Wall Thickness	-0.196	p4352	Axial WRS Pre-mitigation	-0.148
p2592	Comp-to-Comp Variab Factor, fcomp	0.184	p2592	Comp-to-Comp Variab Factor, fcomp	0.084
p4352	Axial WRS Pre-mitigation	-0.170	p1102	Pipe Wall Thickness	-0.070
p5402	Slope, B1 (circ)	0.105	p2595	Charact Width of Peak vs ECP, c	0.060
Probability of Circumferential Rupture – $R^2 = 0.420$			Maximum Axial Outer Half-Length – $R^2 = 0.540$		
p2543	Multiplier proport. Const. A (DM1)	0.535	p2543	Multiplier proport. Const. A (DM1)	0.628
p2592	Comp-to-Comp Variab Factor, fcomp	0.196	p2592	Comp-to-Comp Variab Factor, fcomp	0.224
p1102	Pipe Wall Thickness	-0.192	p4350	Hoop WRS Pre-mitigation	-0.196
p4352	Axial WRS Pre-mitigation	-0.165	p2595	Charact Width of Peak vs ECP, c	0.179
p5402	Slope, B1 (circ)	0.103	p2594	Peak-to-Valley ECP Ratio - 1, P-1	-0.156
Total Leak Rate – $R^2 = 0.566$			Maximum Circumferential Outer Half-Length – $R^2 = 0.420$		
p2543	Multiplier proport. Const. A (DM1)	0.662	p2543	Multiplier proport. Const. A (DM1)	0.538
p2592	Comp-to-Comp Variab Factor, fcomp	0.216	p2592	Comp-to-Comp Variab Factor, fcomp	0.187
p2595	Charact Width of Peak vs ECP, c	0.145	p1102	Pipe Wall Thickness	-0.186
p1102	Pipe Wall Thickness	-0.145	p4352	Axial WRS Pre-mitigation	-0.170
p2594	Peak-to-Valley ECP Ratio - 1, P-1	-0.111	p5402	Slope, B1 (circ)	0.099

Table 6 indicates that variable p2543 is consistently the most significant variable because the magnitude of its SRRC is much higher than all of the other sampled variables used for this scenario. This variable is used in the PWSCC crack initiation model (DM1) in xLPR. After p2543, it is seen that p4350 (ID hoop pre-mitigation WRS) axial results and p4352 (ID axial pre-mitigation WRS) for circumferential results are seen to have high importance as they consistently appear within the top five results.

Similar to previous scenarios (see Section 3.3.2), the 95th quantile was chosen for importance sampling of p2543. This sampling selection led to a larger number of circumferential ruptures, as shown in Figure 94 for Runs 3 and 4. This result confirms both the importance of variable p2543 to the results of this scenario and that the importance sampling generated a greater number of events for a fixed number of samples.

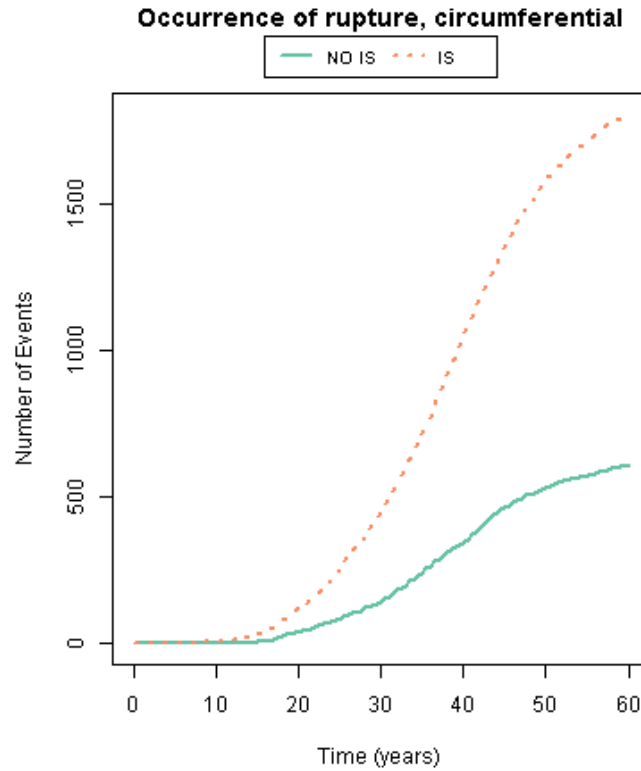


Figure 94: Number of circumferential rupture events occurring for Scenario 5, Run 3 (NO IS) and Run 4 (IS).

Figure 95 and Figure 96 compare Runs 3 and 4 with and without importance sampling for sample size of 1000/50 using LHS. The width of the confidence interval does not decrease at 60 years using importance sampling. A separate study on the importance of the quantile selected for importance sampling is described in Section 4.2.

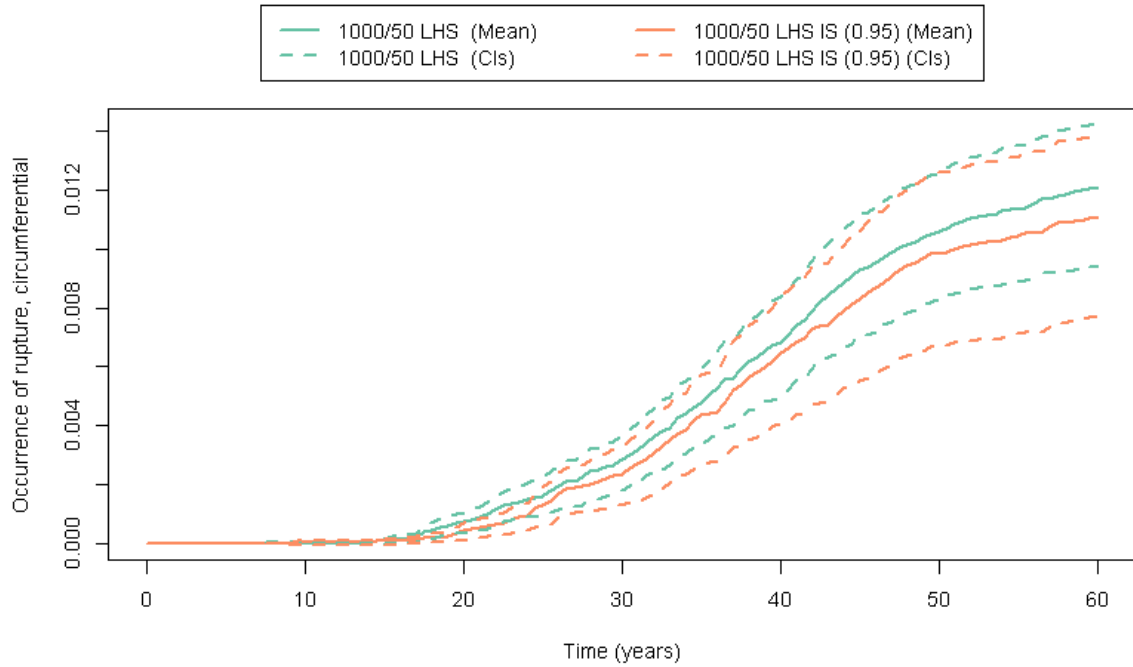


Figure 95: Mean probability of occurrence of circumferential rupture (solid line) and 95% confidence intervals (dashed lines) for Scenario 5, Run 3 (green) and Run 4 (orange).

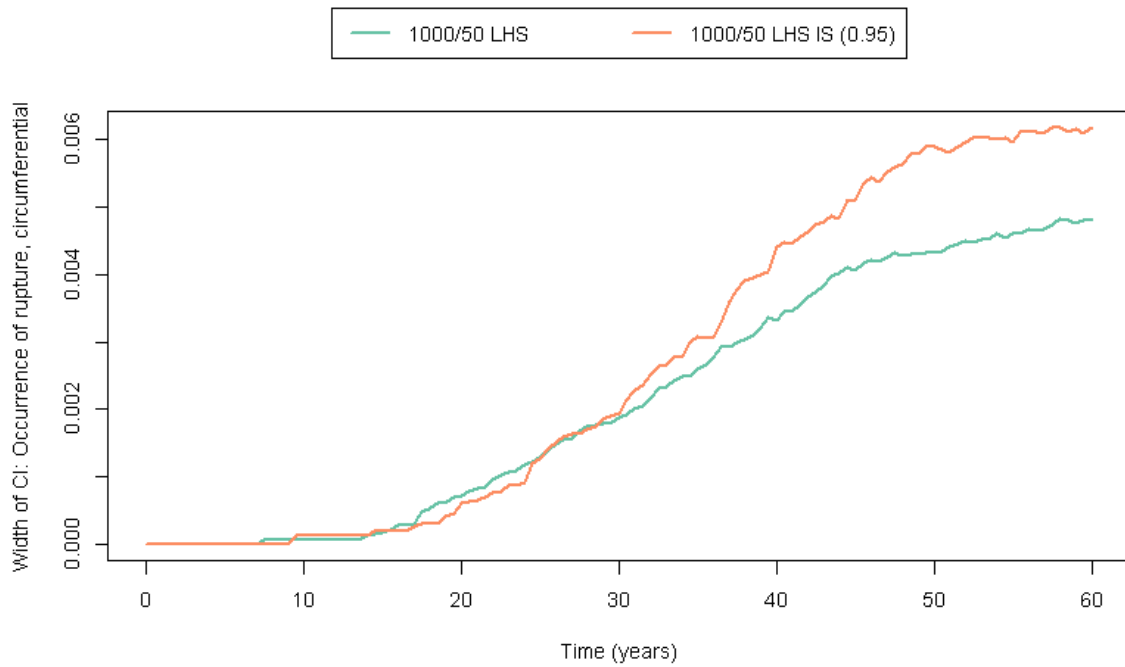


Figure 96: Width of 95% confidence interval for probability of occurrence of axial cracks for Scenario 5, Run 3 (green) and Run 4 (orange).

3.5.3 Overall Results

Figure 97 compares the results for the mean probability of occurrence of circumferential ruptures for Runs 1 through 5. Based on Figure 97, sampling schemes that use SRS with no importance sampling produce results that are significantly different than other sampling schemes. Sampling schemes that use importance sampling and/or LHS (i.e., Runs 3 through 5) are approximately equal with respect to the magnitude of the mean over the simulation time. This is further illustrated in Figure 98 for mean probability of occurrence of circumferential cracks.

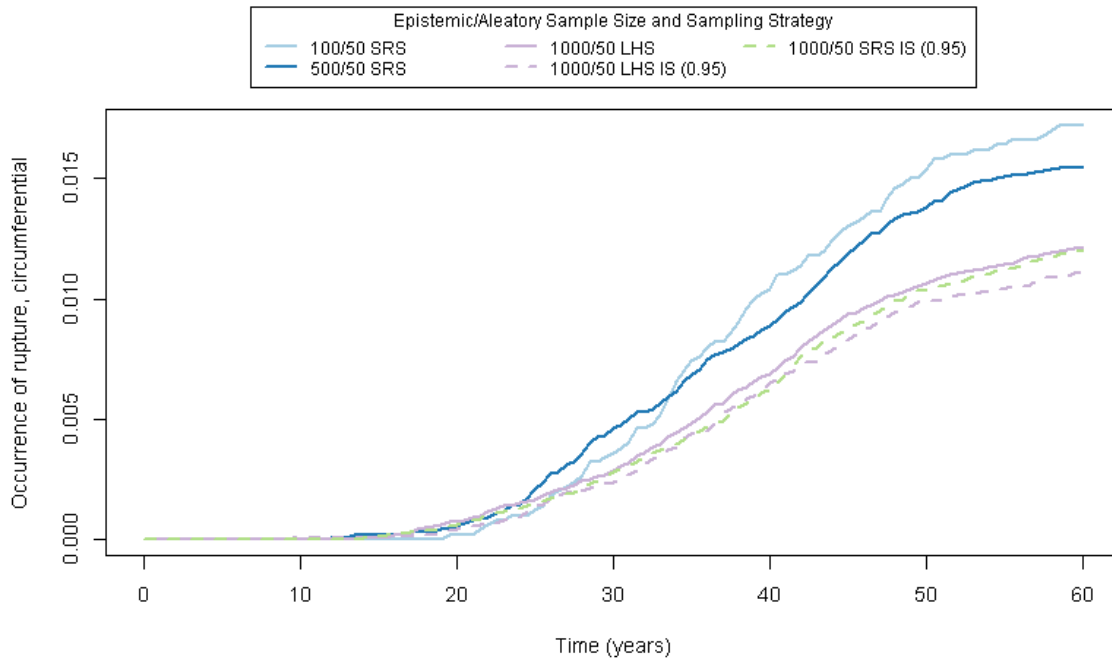


Figure 97: Mean probability of occurrence of circumferential rupture for Scenario 5, Runs 1 through 5.

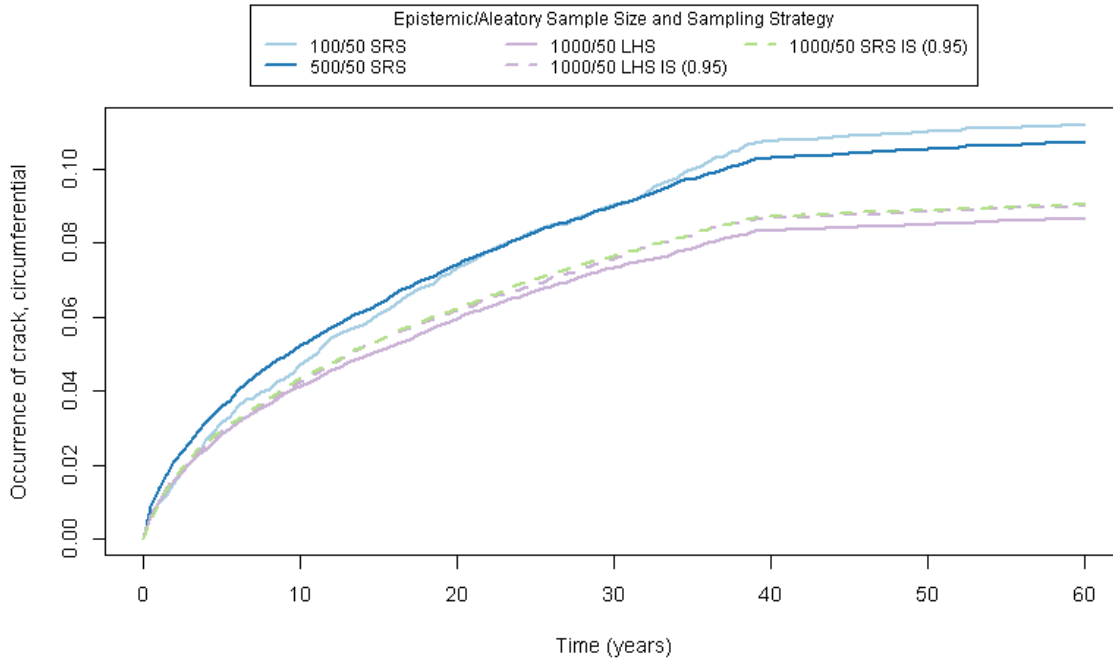


Figure 98: Mean probability of occurrence of circumferential cracks for Scenario 5, Runs 1 through 5.

3.5.4 Epistemic Uncertainty Analysis

An uncertainty analysis was conducted for the sampling scheme used in Run 4 (1000 epistemic and 50 aleatory samples and LHS with IS on p2543) with the goal of characterizing epistemic uncertainty in the probabilities of occurrence.

Plots of the estimated probabilities of occurrence of circumferential crack and rupture are displayed in Figure 99 and Figure 100, respectively. These figures show the probability of occurrence for each epistemic realization (grey lines), the mean (red line), and the 5th, 50th, and 95th percentiles (blue, green, and purple lines, respectively). As noted in Section 2.5.4, the percentiles shown here describe the actual variation found in the particular set of epistemic results.

Figure 99 shows that 5% (95th percentile) of epistemic realizations have an occurrence of circumferential cracks that is greater than 0.44 at year 60. Figure 100 shows that half of the epistemic realizations (i.e., the 50th percentile) have a 0% chance of circumferential rupture occurrence at year 60.

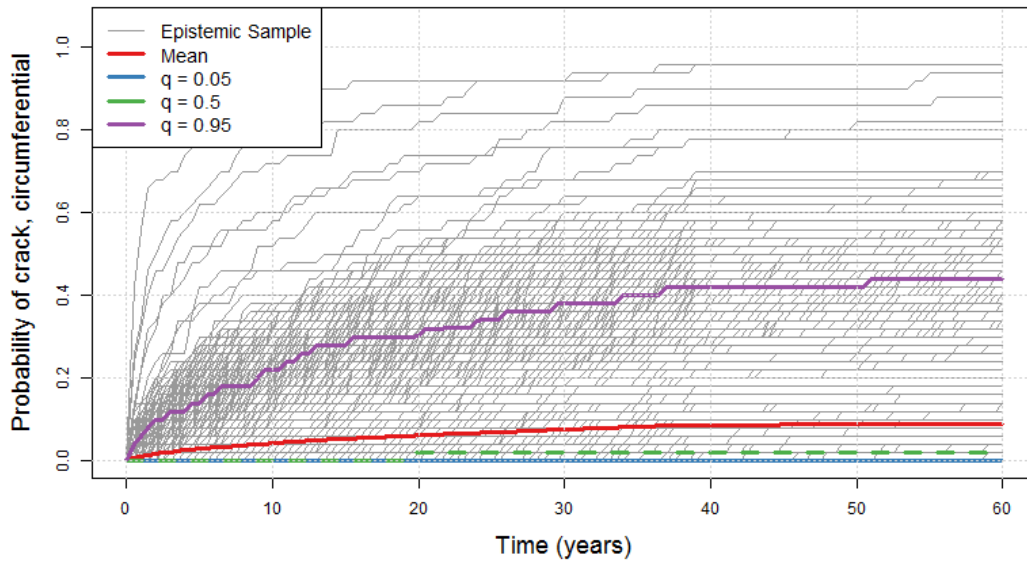


Figure 99: Probability of occurrence of circumferential cracks for each epistemic realization (grey), the mean (red), and the 5th (blue), 50th (green), and 95th (purple) percentiles for Scenario 5, Run 4.

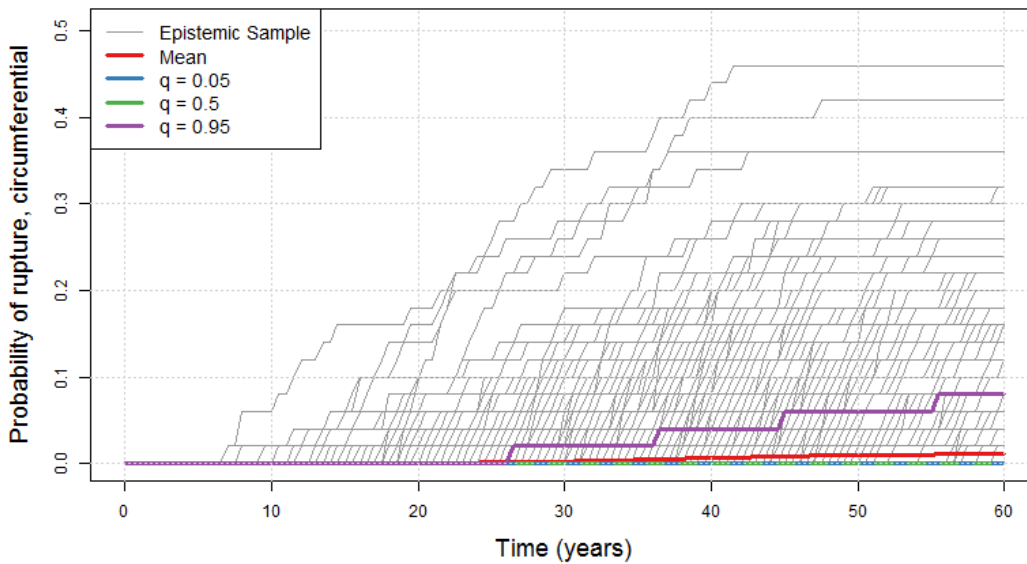


Figure 100: Probability of occurrence of circumferential rupture for each epistemic realization (grey), the mean (red), and the 5th (blue), 50th (green), and 95th (purple) percentiles for Scenario 5, Run 4.

Figure 101 compares the mean and 95th percentile for the probability of occurrence of circumferential leaks and ruptures. As expected, circumferential leaks precede ruptures but

comparing the mean probabilities, circumferential leaks are quickly followed by ruptures. The delay between occurrence of leaks and ruptures are greater for the 95th percentile but at year 60, both events have equal probabilities of occurrence.

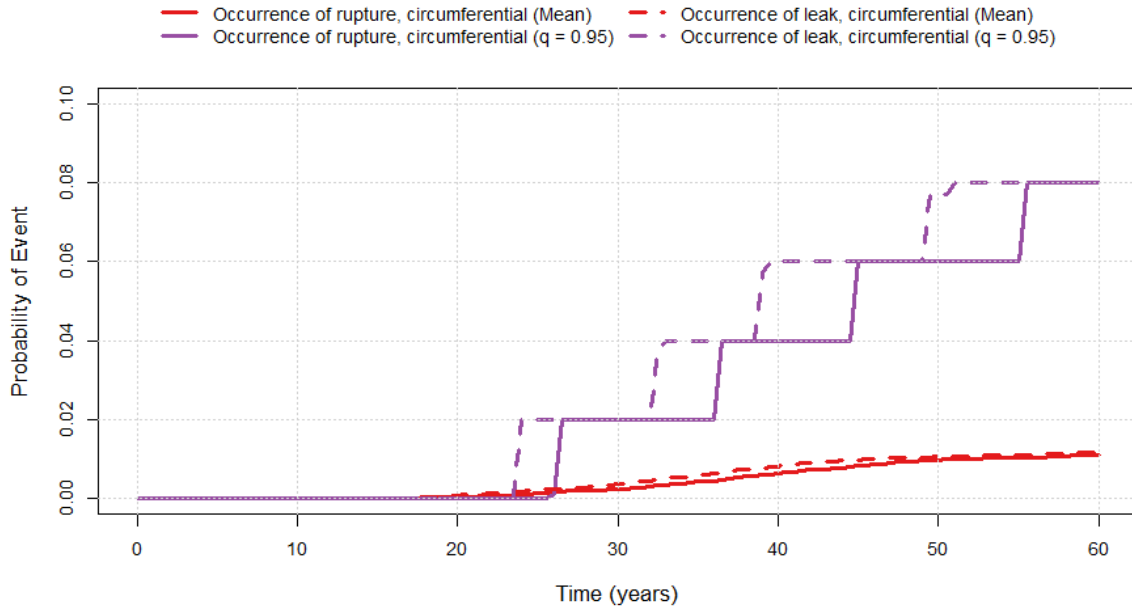


Figure 101: Mean (red) and 95th percentile (purple) for occurrence of circumferential leak (dashed) and rupture (solid) for Scenario 5, Run 4.

3.5.5 Convergence Analysis

The convergence of the sampling scheme for Scenario 4, Run 4 (1000 epistemic samples with LHS and IS) was analyzed using five replicate simulations with different random number seeds for both the epistemic and aleatory loops. Figure 102 shows the mean probability of occurrence of circumferential cracks for each replicate and Figure 103 shows the mean probability of occurrence of circumferential rupture for each replicate. A 95% prediction interval was created around the overall mean of the five replicates, as seen in Figure 104 and Figure 105 for occurrence of circumferential crack and rupture, respectively. The width of the 95% prediction interval is relatively narrow over the course of the simulation time for occurrence of circumferential cracks. For occurrence of circumferential rupture, the width of the 95% prediction interval is about 1% at 60 years.

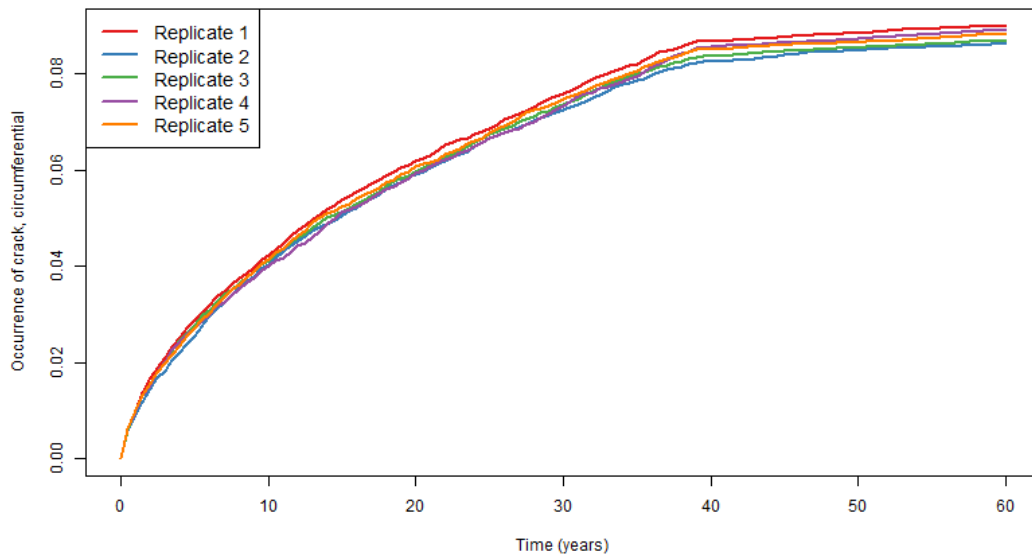


Figure 102: Comparison of estimated mean probability of occurrence of circumferential cracks for five convergence replicates for Scenario 5, Run 4.

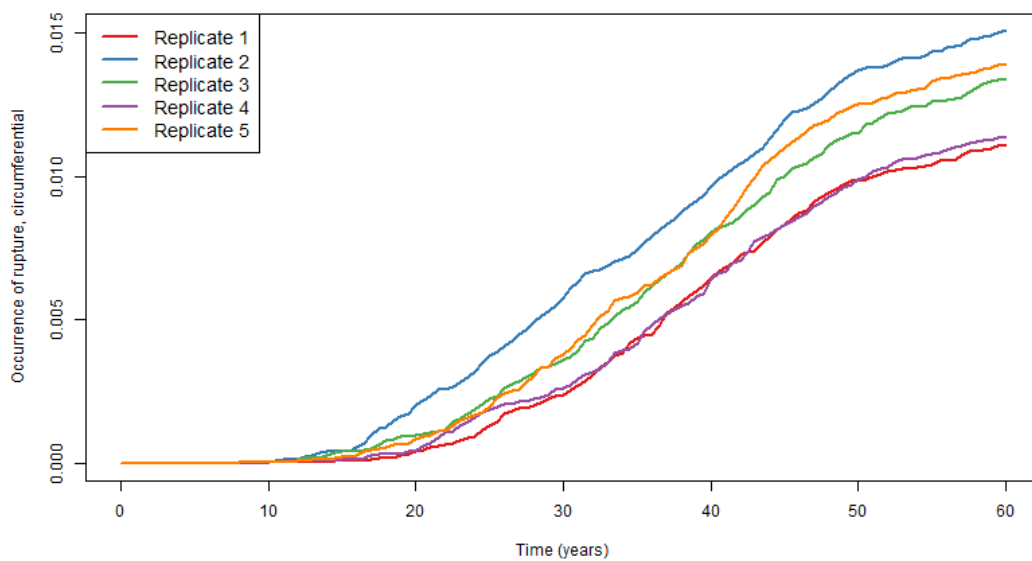


Figure 103: Comparison of estimated mean probability of occurrence of circumferential rupture for five convergence replicates for Scenario 5, Run 4.

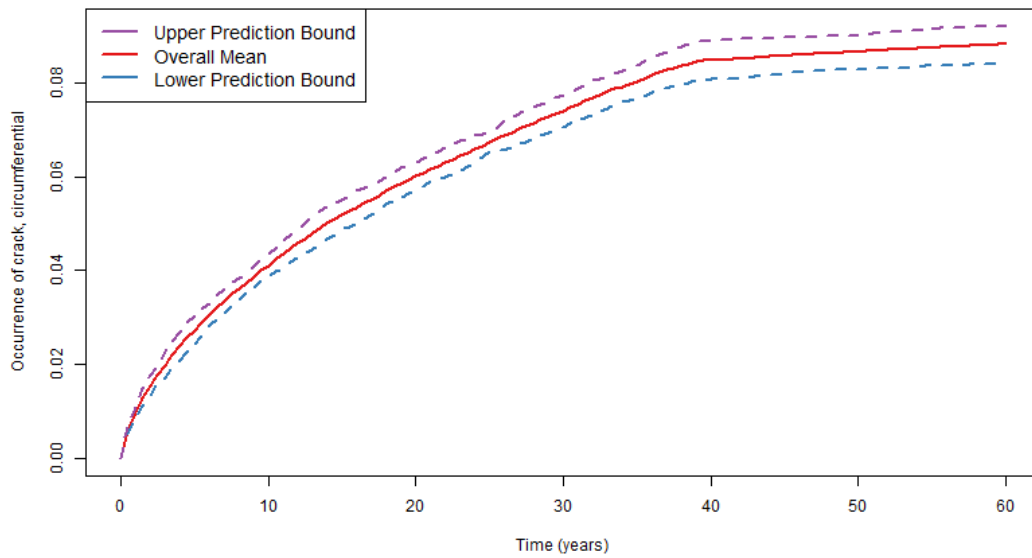


Figure 104: 95% prediction interval over mean probability of circumferential cracks for five replicates for Scenario 5, Run 4.

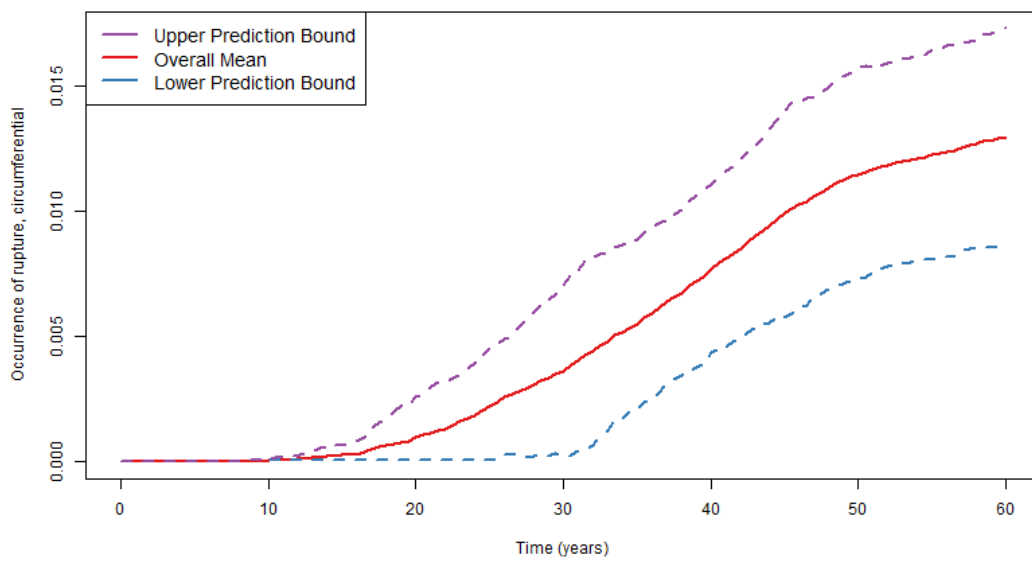


Figure 105: 95% prediction interval over mean probability of circumferential rupture for five replicates for Scenario 5, Run 4.

3.5.6 Comparison to Scenario 3 Results

Scenario 5 runs were compared to Scenario 3 runs with corresponding sampling selections. To properly compare the two scenarios, Run 9 from Scenario 3 and Run 4 from Scenario 5 were used for the comparison as both runs represent a run with 1000 epistemic samples with LHS and IS.

Figure 106 shows the effects MSIP mitigation at 40 years on the probability of occurrence of axial and circumferential cracks. The trends that were observed for axial and circumferential cracks in Scenario 4 are also observed for Scenario 5. That is, after stress mitigation is implemented, formation of axial cracks goes unchanged for the rest of the simulation while circumferential cracks continue to form, but at a significantly slower rate compared to Scenario 3. The post-mitigation compressive axial and hoop stress is reduced, but in the case of hoop stress, the compressive stress is reduced enough such that no more axial cracks form. Post-mitigation axial stresses are not reduced enough to completely stop circumferential crack formation, but as seen in the following results, the rate of crack formation and growth is reduced.

The general trend observed for probability of occurrence of crack results is also observed for occurrence of axial and circumferential leak results as seen in Figure 107. Figure 108 provides a comparison for occurrence of circumferential rupture (note that axial rupture is not predicted for Scenario 3 or Scenario 5). In the case of circumferential cracks, occurrence of leak and rupture are strongly correlated, as expected.

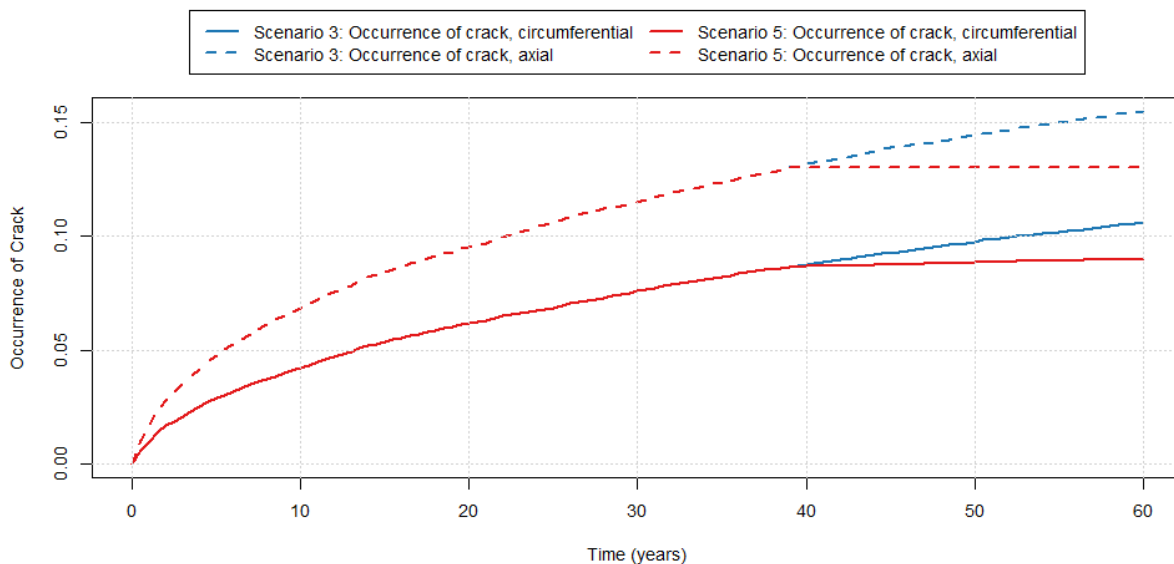


Figure 106: Mean probability of axial (dashed lines) and circumferential (solid lines) crack occurrence for Scenario 3 (blue lines) and Scenario 5 (red lines).

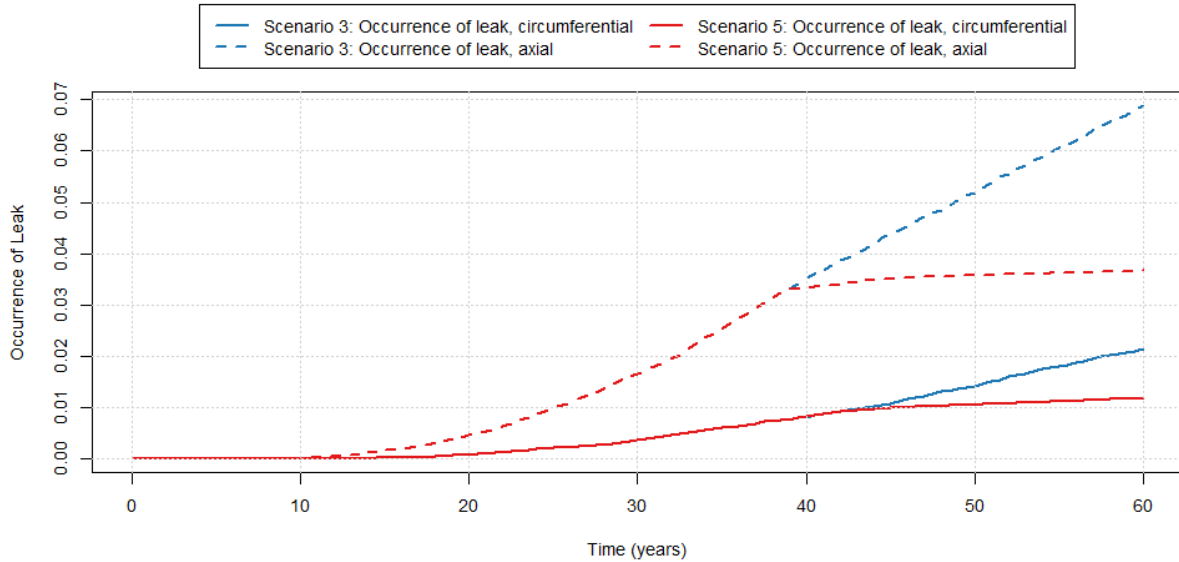


Figure 107: Mean probability of axial (dashed lines) and circumferential (solid lines) leak occurrence for Scenario 3 (blue lines) and Scenario 5 (red lines).

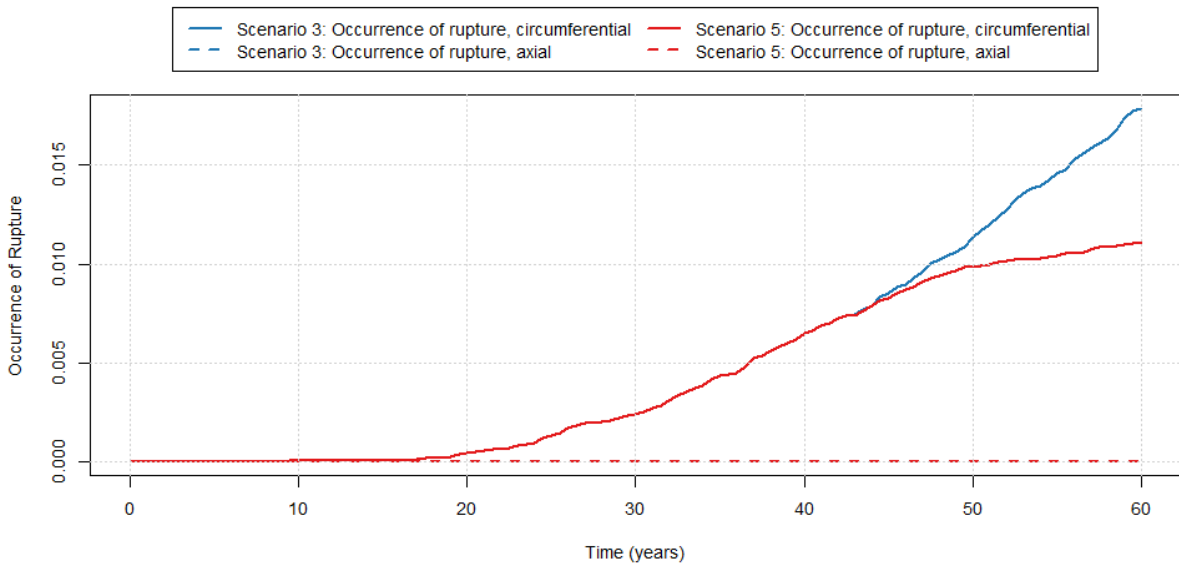


Figure 108: Mean probability of axial (dashed lines) and circumferential (solid lines) rupture occurrence for Scenario 3 (blue lines) and Scenario 5 (red lines).

Scenario 4 and Scenario 5 have identical post-mitigation WRS profiles, but the mitigation occurs at different times: 20 years and 40 years for Scenario 4 and Scenario 5, respectively. Scenario 4 and Scenario 5 comparisons for occurrence of crack, occurrence of leak, occurrence of rupture, and total leak rate are plotted in Figure 109, Figure 110, Figure 111, and Figure 112, respectively. Occurrence of leaks (both axial and circumferential), occurrence of circumferential rupture, and total leak rate increase about six times more when MSIP mitigation is implemented at 40 years versus 20 years.

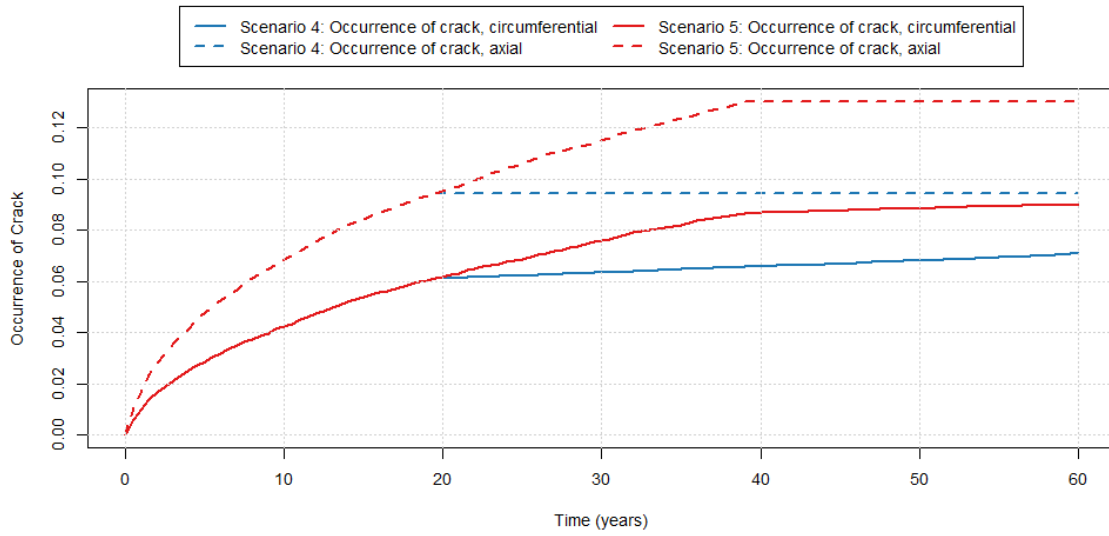


Figure 109: Mean probability of axial (dashed lines) and circumferential (solid lines) crack occurrence for Scenario 4 (blue lines) and Scenario 5 (red lines).

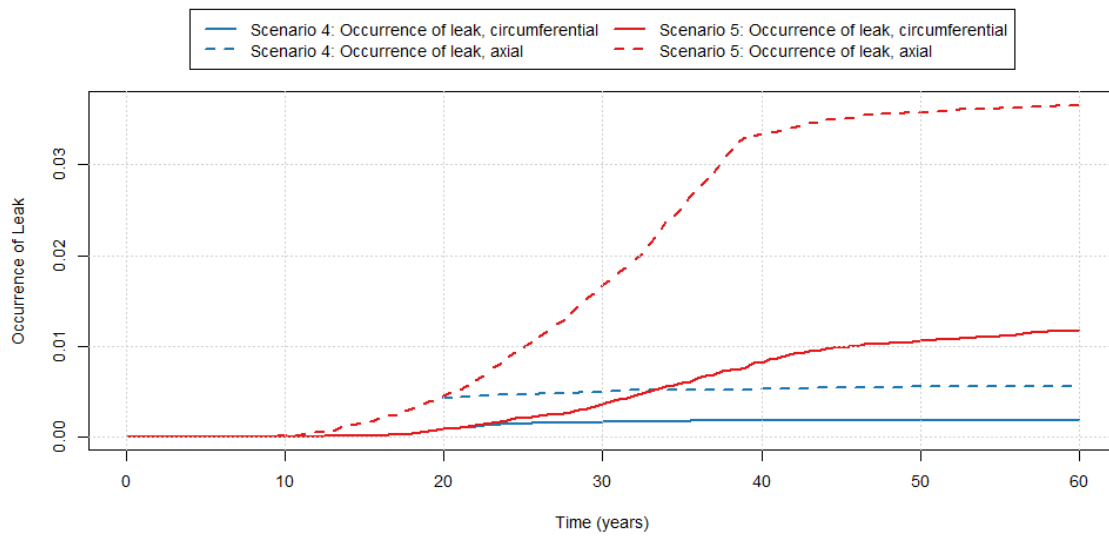


Figure 110: Mean probability of axial (dashed lines) and circumferential (solid lines) leak occurrence for Scenario 4 (blue lines) and Scenario 5 (red lines).

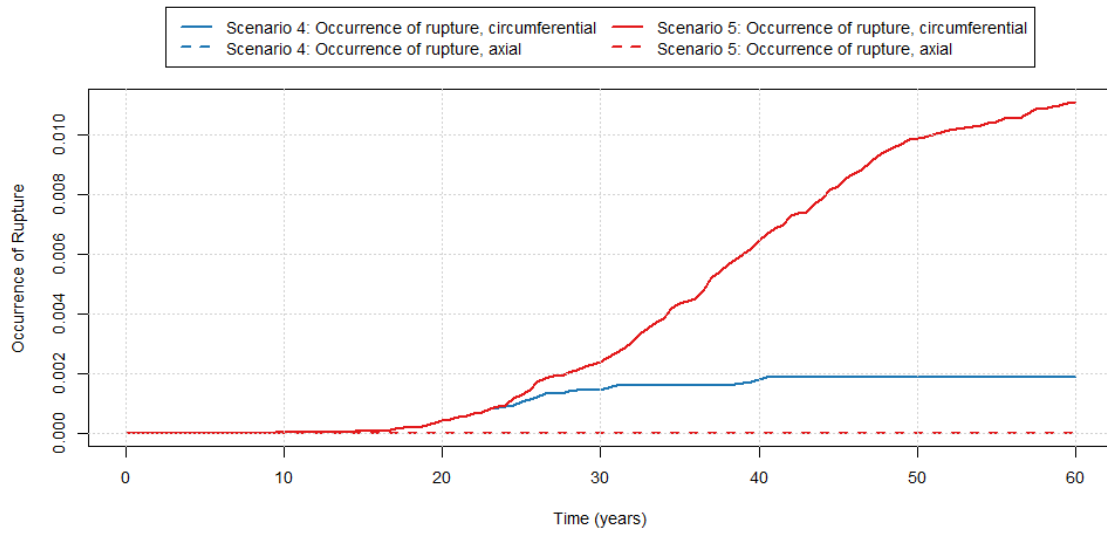


Figure 111: Mean probability of axial (dashed lines) and circumferential (solid lines) rupture occurrence for Scenario 4 (blue lines) and Scenario 5 (red lines).

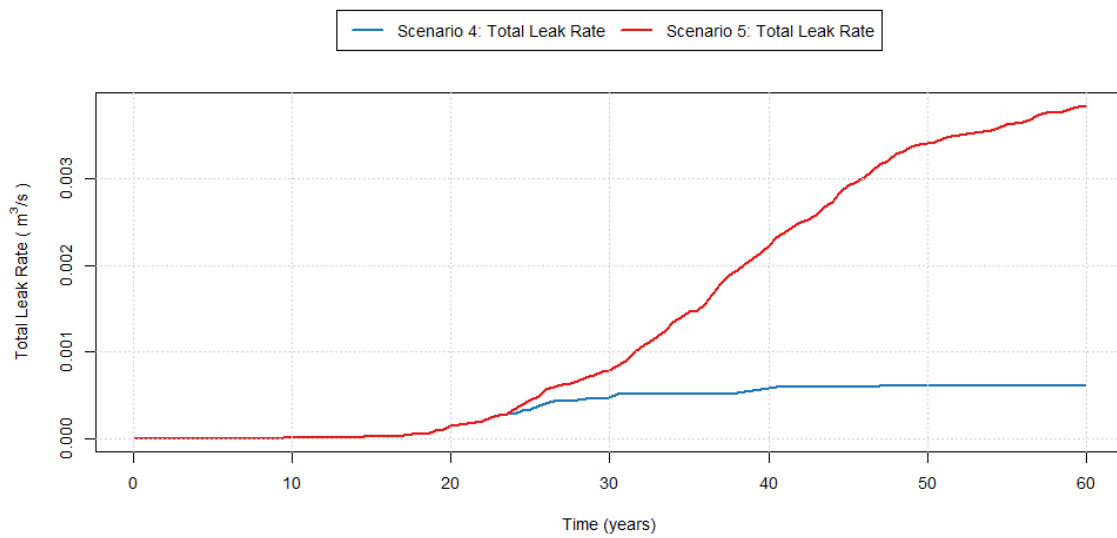


Figure 112: Mean total leak rate for Scenario 4 (blue line) and Scenario 5 (red line).

3.6 Scenario 6

3.6.1 Scenario Summary

The results and conclusions of a detailed analysis of sampling schemes and their impact on an understanding of the uncertainty found in the simulation results of Scenario 6 are given in this section.

Scenario 6 uses the following defining options in the xLPR code:

- PWSCC flaw initiation
- Circumferential and axial flaw orientation
- PWSCC flaw growth
- Chemical mitigation Zinc (Zn) at 20 years

This scenario is based on Scenario 3, but includes chemical mitigation with Zn at 20 years.

The outputs of interest defined for this scenario for both circumferential and axial cracks are:

- Probability of occurrence of crack
- Probability of occurrence of leak
- Probability of occurrence of rupture
- Total Leak Rate
- Maximum crack depth
- Maximum crack inner half-length
- Maximum crack outer half-length

The comprehensive list of simulations that were used in the analysis of this scenario is shown below. A constant aleatory sample size of 50 was used for each of the sampling schemes considered.

1. 100/50 (epistemic/aleatory sample size) with simple random sampling (SRS)
2. 500/50 with simple random sampling (SRS)
3. 1000/50 with Latin hypercube sampling (LHS)
4. 1000/50 with LHS and IS on p2543 with target quantile of 0.95
5. 1000/50 with SRS and IS on p2543 with target quantile of 0.95

Sampling schemes were chosen based on insight gained from the results of the Scenario 3 analysis; those schemes that were most informative were chosen for Scenario 6.

The results of the Scenario 6 analysis are consistent with the analysis of Scenario 3. The only difference between the scenarios is that Scenario 6 includes Zn mitigation at 20 years; in every other respect the scenario is identical to Scenario 3. While many outputs of interest were defined for this scenario, the results of regression analyses indicate that the same variable, the multiplier on Direct Model 1 (DM1) proportionality constant A for the PWSCC crack initiation module, hereafter referred to as p2543, was the most important input for explaining variability in all of the outputs for Scenario 6. For the sake of brevity, the only result shown in this section is the occurrence of circumferential rupture, as the results for other outputs are comparable with respect to convergence

and stability. When used, importance sampling was applied to p2543 to improve the precision in the output of interest.

As in the Scenario 3 analysis, overall, results confirm that increasing the epistemic sample size decreases the uncertainty in the mean of the output (e.g., probability of occurrence of leak from axial cracks) as expected, especially when moving from 100 to 500 epistemic samples. The decrease in uncertainty in the mean of the output is also evident when transitioning from 500 to 1000 epistemic realizations. An epistemic sample size of 1000 is sufficient for this scenario, though a sample of 500 may also provide sufficient mean precision depending on user needs.

3.6.2 Analysis Progression

The five sampling schemes, informed by the Scenario 3 analysis in Section 3.3, were chosen to demonstrate stability and the impact of sampling options on the stability of results. These schemes progress in sample size and sampling options.

Runs 1 and 2 are identical in all sampling options except the epistemic sample size. The epistemic sample size for Run 1 was 100 and this was increased to 500 for Run 2. The mean probability of occurrence of axial cracks from Runs 1 and 2 is plotted in Figure 113 with 95% confidence intervals calculated using the bootstrap method. The mean estimates for both runs deviate at about 20 years and diverge more from 55 to 60 years. This is similar to Scenario 3, showing no substantive difference in stability due to mitigation.

With regard to stability, the width of the confidence intervals around the estimated mean decreases as expected when the epistemic sample size is increased to 500. This is shown in Figure 114. Whether this decrease is significant is application specific; with 100 epistemic samples, the estimated probability of occurrence of circumferential ruptures at year 60 ranges from about 0.14 to 0.39. Increasing the sample size to 500 reduces the width of this interval from about 0.25 to about 0.01.

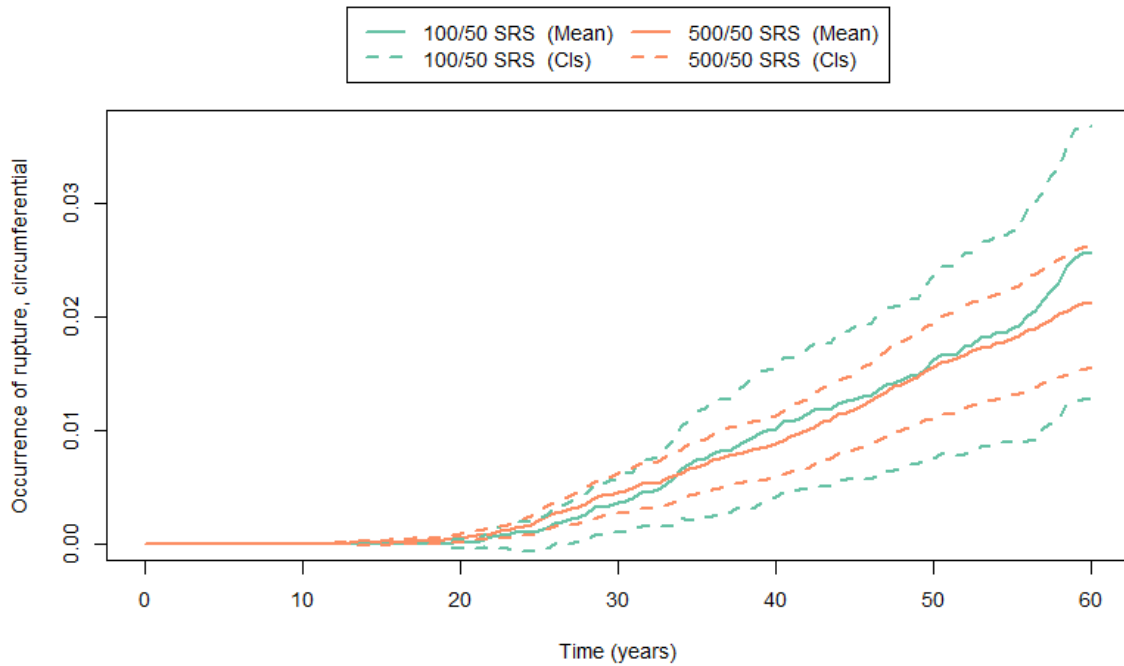


Figure 113: Mean probability of occurrence of circumferential ruptures (solid lines) and 95% confidence intervals (dashed lines) for Scenario 6, Run 1 (green) and Run 2 (orange).

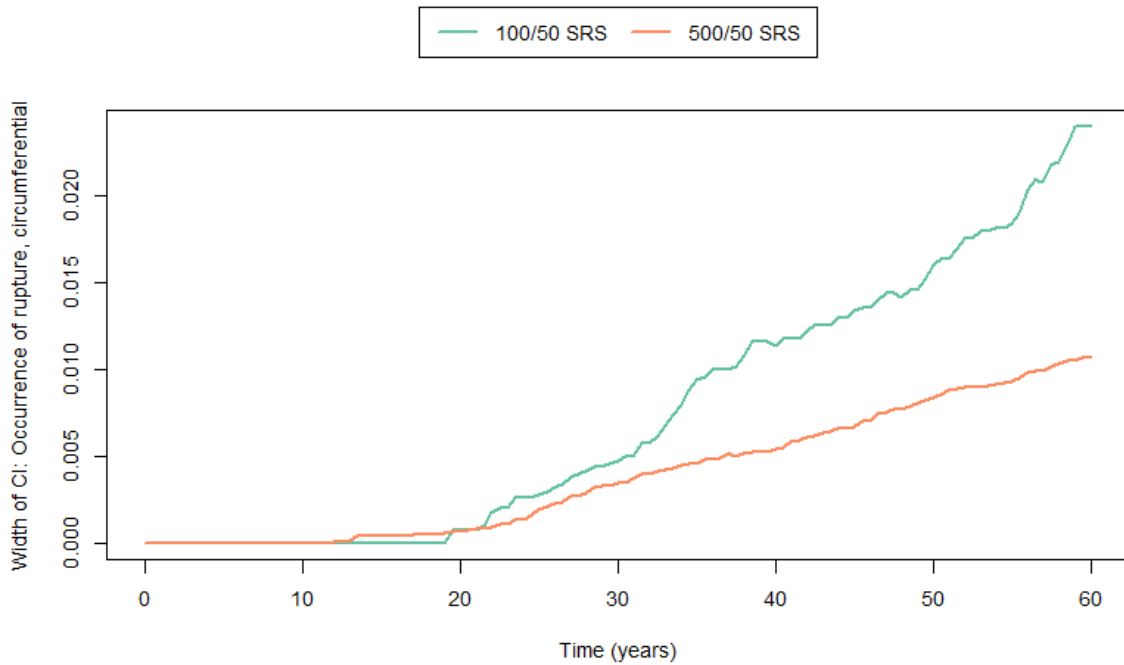


Figure 114: Width of 95% confidence interval for the estimate of mean probability of occurrence of circumferential ruptures for Scenario 6, Run 1 (green) and Run 2 (orange).

The relatively large decrease in confidence interval width between Runs 1 and 2 suggests that an increase in the number of samples, and thus simulation time, may still be worth the benefit in reducing uncertainty in the final mean estimate. The Scenario 3 analysis showed that LHS sampling could improve stability and there is no computational reason to prefer SRS to LHS. Thus, to improve upon the results in Run 2, the epistemic sample size was increased to 1000 and LHS was used instead of SRS to sample epistemically uncertain parameters for Scenario 6.

The increase in sample size and change in sampling strategy from Run 2 to Run 3 decreased the width of the confidence interval by a much smaller margin, shown in Figure 115 and Figure 116. However, the mean probability of occurrence of axial cracks is substantially lower when LHS is employed compared to SRS; the mean estimate from Run 2 is not contained in the confidence intervals around the mean estimate from Run 3. This motivates the inclusion of IS to examine stability.

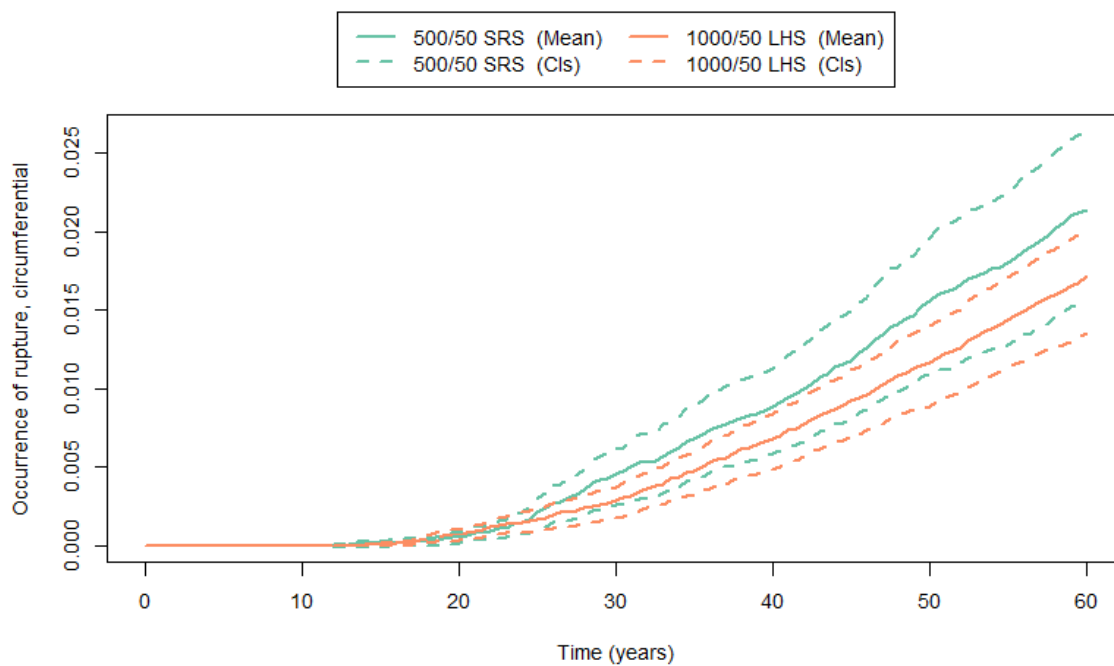


Figure 115: Mean probability of occurrence of circumferential ruptures (solid lines) and 95% confidence intervals (dashed lines) for Scenario 6, Run 2 (green) and Run 3 (orange).

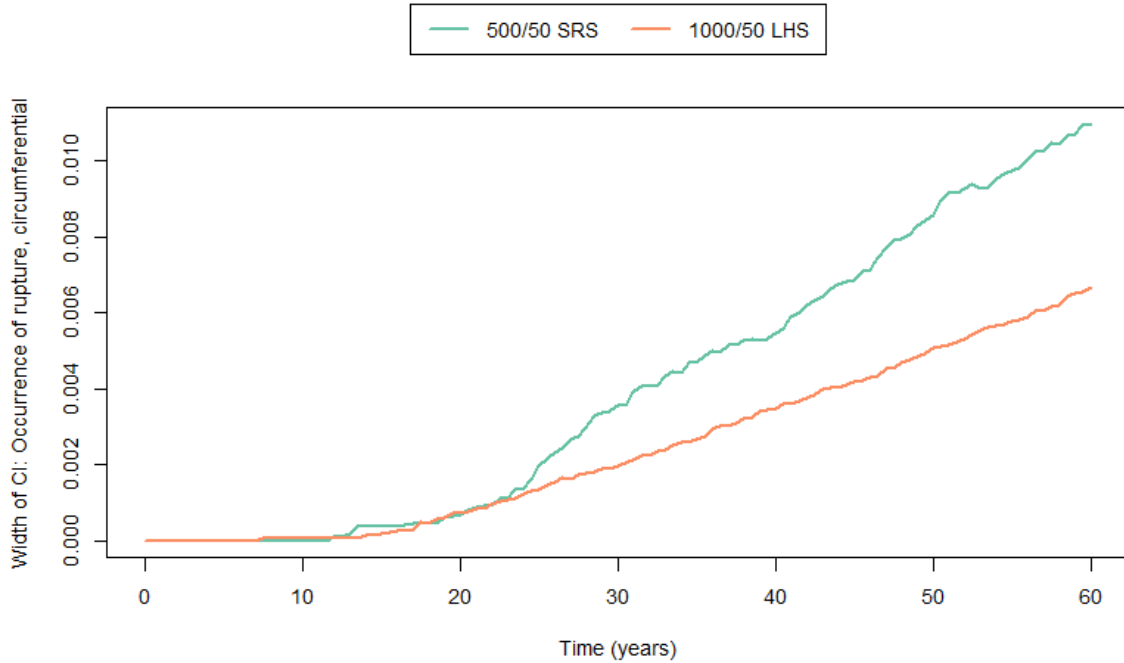


Figure 116: Width of the 95% confidence interval for probability of occurrence of circumferential ruptures for Scenario 6, Run 2 (green) and Run 3 (orange).

Recall that importance sampling concentrates sampling more in regions of the input space that are expected to contribute most to variation in the simulation results. Stepwise rank regression results from Run 3 on the probability of occurrence of circumferential ruptures for Scenario 6 are presented in Table 7. These results are consistent with regression results from Runs 1 and 2. The parameter identified in Runs 1, 2, and 3 Scenario 6 as the most important variable with respect to variability in the probability of occurrence of circumferential ruptures is p2543, the multiplier on DM1 proportionality constant A. The influence of this variable can be seen both by the high SRRC value (0.595) in the table and the scatterplots of p2543 with the probability of occurrence of circumferential ruptures in Figure 117. The increasing occurrence of circumferential ruptures as p2543 increases is consistent in all three runs.

Table 7: Summary of the stepwise rank regression analysis results for all outputs at year 60 for Scenario 6, Run 3.

Variable Identifier	Variable Name	SRRC	Variable Identifier	Variable Name	SRRC
Probability of Occurrence of Axial Cracks – $R^2 = 0.862$			Maximum Axial Crack Depth – $R^2 = 0.809$		
p2543	Multiplier proport. Const. A (DM1)	0.909	p2543	Multiplier proport. Const. A (DM1)	0.855
p4350	Hoop WRS Pre-mitigation	-0.186	p4350	Hoop WRS Pre-mitigation	-0.234
p5106	b (circ)	0.065	p1102	Pipe Wall Thickness	-0.083
p2531	Zn Factor of Improvement - 1, FOIZn-1	-0.056	p2592	Comp-to-Comp Variab Factor, fcomp	0.075
p1102	Pipe Wall Thickness	-0.049	p2595	Charact Width of Peak vs ECP, c	0.059
Probability of Occurrence of Circumferential Cracks – $R^2 = 0.809$			Maximum Circumferential Crack Depth – $R^2 = 0.786$		
p2543	Multiplier proport. Const. A (DM1)	0.891	p2543	Multiplier proport. Const. A (DM1)	0.863
p4352	Axial WRS Pre-mitigation	-0.116	p4352	Axial WRS Pre-mitigation	-0.154
p2531	Zn Factor of Improvement - 1, FOIZn-1	-0.071	p2592	Comp-to-Comp Variab Factor, fcomp	0.066
p1102	Pipe Wall Thickness	-0.022	p1102	Pipe Wall Thickness	-0.061
p5108	b (axial)	0.021	p2594	Peak-to-Valley ECP Ratio - 1, P-1	-0.048
Probability of Axial Leak – $R^2 = 0.670$			Maximum Axial Inner Half-Length – $R^2 = 0.852$		
p2543	Multiplier proport. Const. A (DM1)	0.739	p2543	Multiplier proport. Const. A (DM1)	0.898
p4350	Hoop WRS Pre-mitigation	-0.235	p4350	Hoop WRS Pre-mitigation	-0.214
p2592	Comp-to-Comp Variab Factor, fcomp	0.192	p5106	b (circ)	0.067
p2595	Charact Width of Peak vs ECP, c	0.130	p1102	Pipe Wall Thickness	-0.055
p1102	Pipe Wall Thickness	-0.096	p2531	Zn Factor of Improvement - 1, FOIZn-1	-0.055
Probability of Circumferential Leak – $R^2 = 0.525$			Maximum Circumferential Inner Half-Length – $R^2 = 0.767$		
p2543	Multiplier proport. Const. A (DM1)	0.641	p2543	Multiplier proport. Const. A (DM1)	0.847
p4352	Axial WRS Pre-mitigation	-0.187	p4352	Axial WRS Pre-mitigation	-0.160
p2592	Comp-to-Comp Variab Factor, fcomp	0.166	p2592	Comp-to-Comp Variab Factor, fcomp	0.092
p1102	Pipe Wall Thickness	-0.152	p2595	Charact Width of Peak vs ECP, c	0.060
p2594	Peak-to-Valley ECP Ratio - 1, P-1	-0.107	p2594	Peak-to-Valley ECP Ratio - 1, P-1	-0.059
Probability of Circumferential Rupture – $R^2 = 0.477$			Maximum Axial Outer Half-Length – $R^2 = 0.669$		
p2543	Multiplier proport. Const. A (DM1)	0.595	p2543	Multiplier proport. Const. A (DM1)	0.738
p2592	Comp-to-Comp Variab Factor, fcomp	0.192	p4350	Hoop WRS Pre-mitigation	-0.236
p4352	Axial WRS Pre-mitigation	-0.175	p2592	Comp-to-Comp Variab Factor, fcomp	0.193
p1102	Pipe Wall Thickness	-0.162	p2595	Charact Width of Peak vs ECP, c	0.131
p2594	Peak-to-Valley ECP Ratio - 1, P-1	-0.107	p1102	Pipe Wall Thickness	-0.095
Total Leak Rate – $R^2 = 0.673$			Maximum Circumferential Outer Half-Length – $R^2 = 0.521$		
p2543	Multiplier proport. Const. A (DM1)	0.758	p2543	Multiplier proport. Const. A (DM1)	0.637
p2592	Comp-to-Comp Variab Factor, fcomp	0.193	p4352	Axial WRS Pre-mitigation	-0.187
p2595	Charact Width of Peak vs ECP, c	0.132	p2592	Comp-to-Comp Variab Factor, fcomp	0.173
p1102	Pipe Wall Thickness	-0.125	p1102	Pipe Wall Thickness	-0.143
p2594	Peak-to-Valley ECP Ratio - 1, P-1	-0.115	p2594	Peak-to-Valley ECP Ratio - 1, P-1	-0.111

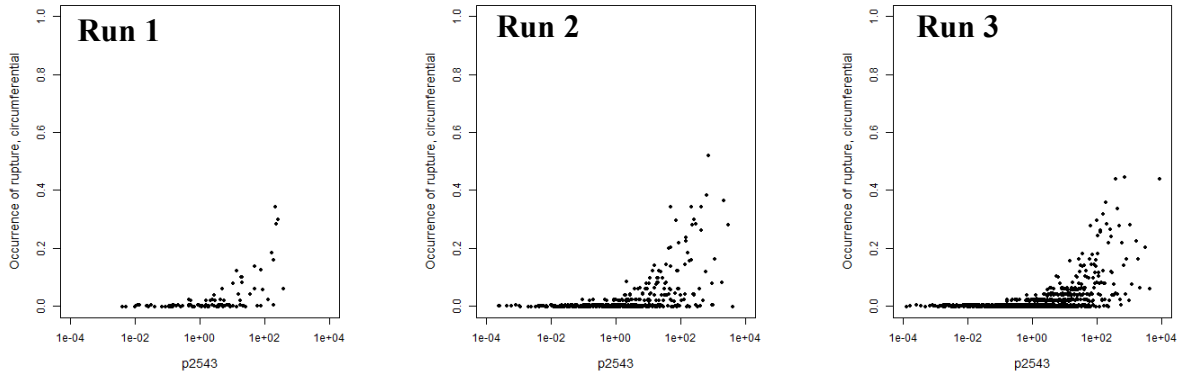


Figure 117: Scatter plots for variable p2543 (multiplier on DM1 proportionality constant A) and the probability of occurrence of circumferential ruptures for Scenario 6, Run 1 (left), Run 2 (center), and Run 3 (right).

Run 4 expands on Run 3 by incorporating importance sampling, while maintaining the same sample size (1000) and sampling strategy (LHS). Importance sampling was applied on p2543 at the 95th quantile to oversample the upper tail of the p2543 distribution to increase the occurrence of

circumferential ruptures. The comparison between number of circumferential rupture events for Runs 3 and 4 is plotted in Figure 118. This plot shows the impact of importance sampling; the run with importance sampling experienced two and three times as many circumferential rupture events as the run without importance sampling. Similarly, the comparison of mean estimates of the probability of occurrence of circumferential ruptures for Runs 3 and 4 is plotted in Figure 119 and the difference between the widths of these confidence intervals around the estimates is plotted in Figure 120. These figures indicate an increase in the number of circumferential ruptures but no decrease in the uncertainty around the estimated mean probability of occurrence of circumferential ruptures from including importance sampling. As discussed in 3.3.2, the impact of importance sampling is dependent upon the quantile chosen to define the sampling and the magnitude of the probability being approximated. That importance sampling did not change the estimate or confidence intervals much may suggest that either a different quantile should be chosen or the probability of occurrence of circumferential ruptures for this scenario is high enough that importance sampling does not provide a benefit. This is evident in the first 25 years of simulation, during which the confidence bounds are narrower when importance sampling is used to estimate the probability of occurrence of rupture. On this time interval, the probability is fairly low and importance sampling seems to provide a benefit that decreases as the estimated probability increases.

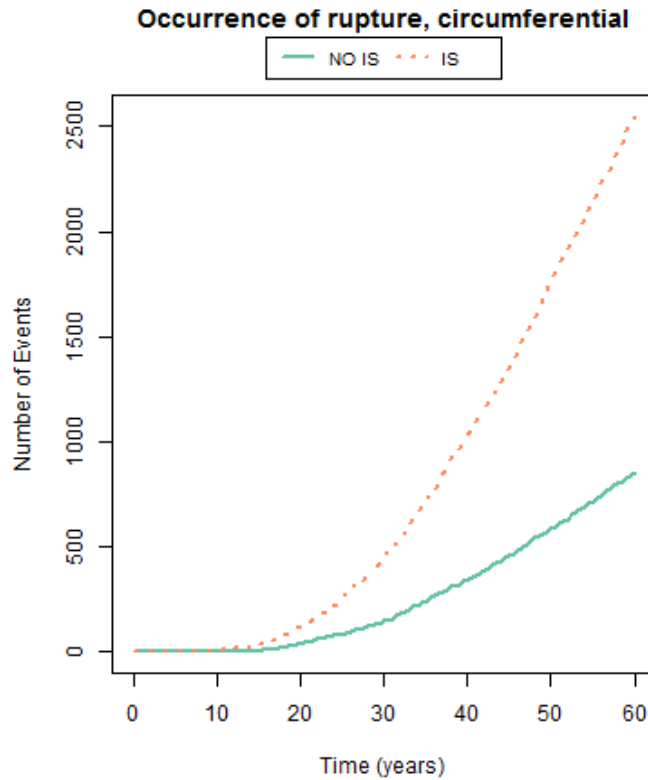


Figure 118: Number of circumferential rupture events occurring for Scenario 6, Run 3 (No IS) and Run 4 (IS).

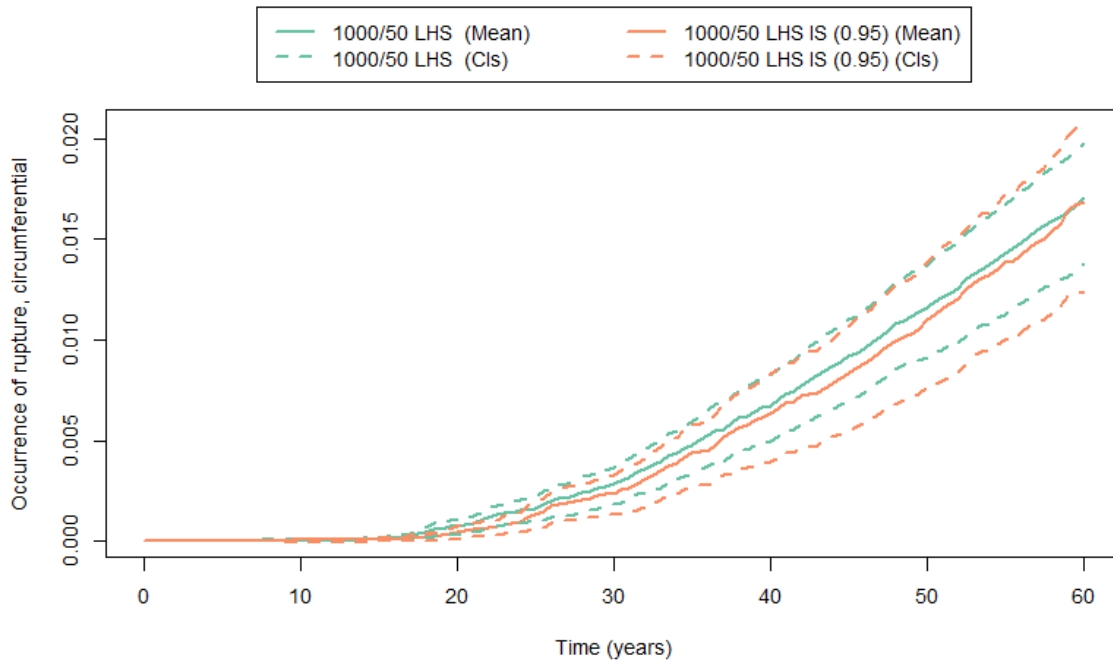


Figure 119: Mean probability of occurrence of circumferential ruptures (solid line) and 95% confidence intervals (dashed lines) for Scenario 6, Run 3 (green) and Run 4 (orange).

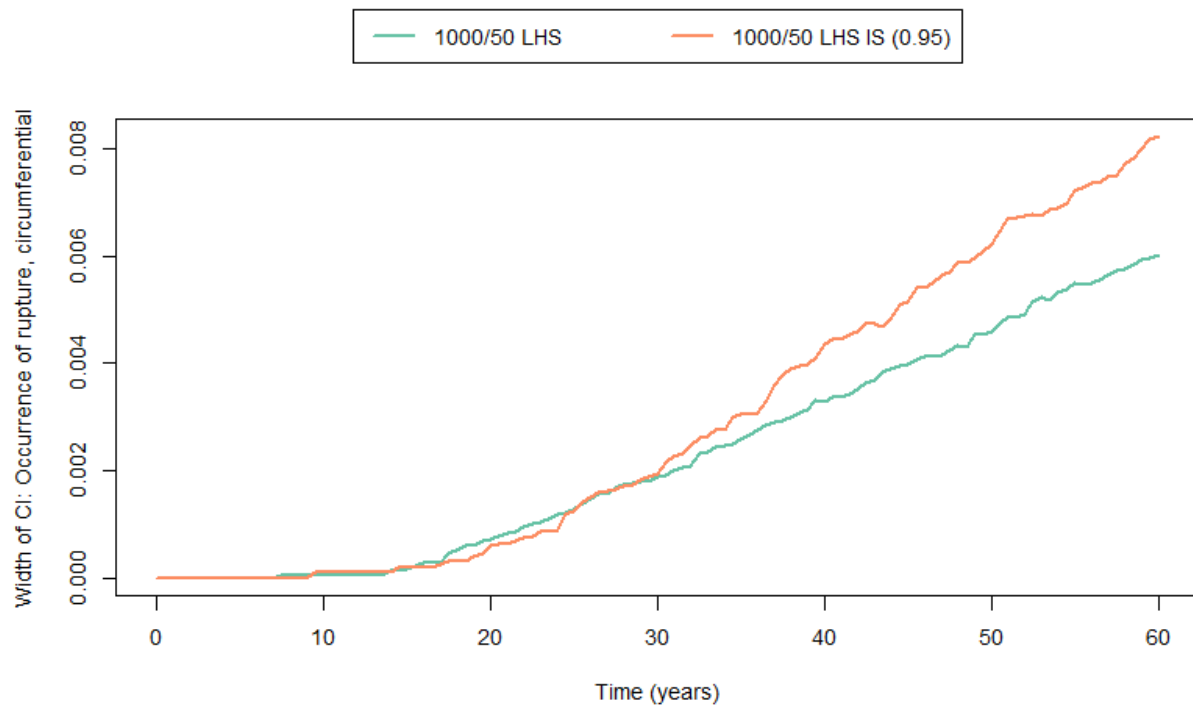


Figure 120: Width of 95% confidence interval for probability of occurrence of circumferential ruptures for Scenario 6, Run 3 (green) and Run 4 (orange).

Stability is further examined by Run 5, which maintains the sample size and importance sampling on p2543 from Run 4, but employs SRS rather than LHS. Even with SRS, the results are very close to the results from Run 4. The estimated mean occurrence of circumferential rupture with confidence intervals from Runs 4 and 5 are plotted in Figure 121 to show the similarities.

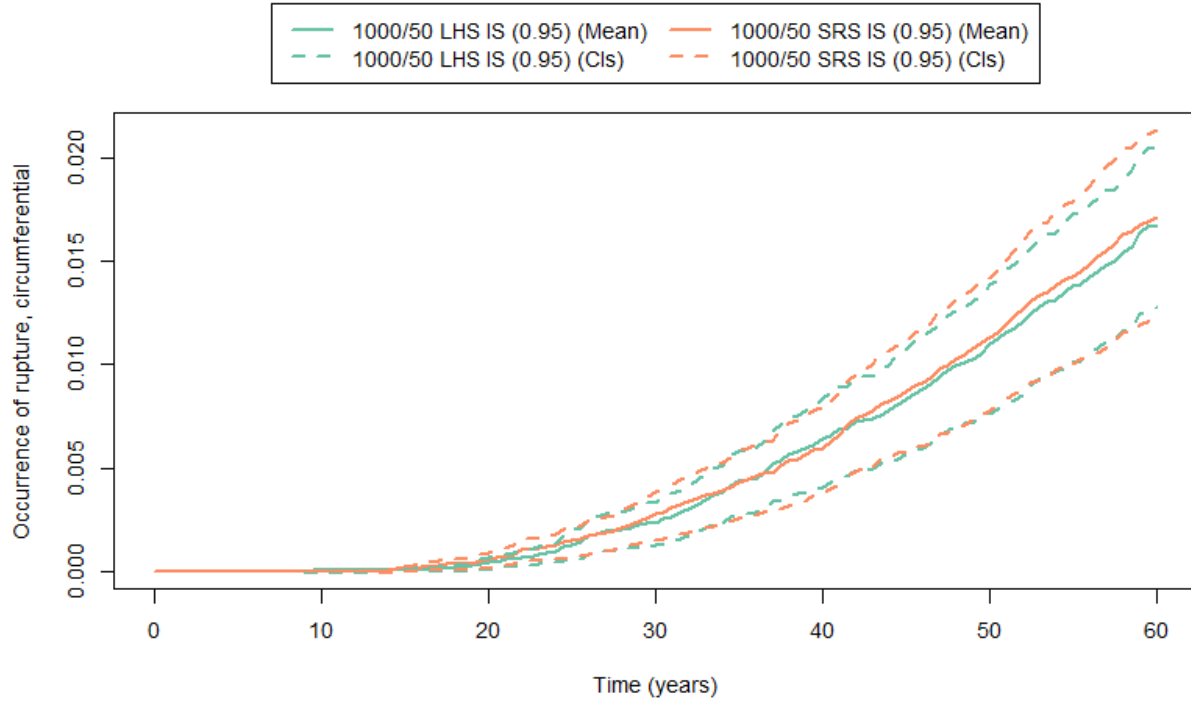


Figure 121: Mean Probability of occurrence of circumferential rupture (solid line) and 95% confidence intervals (dashed lines) for Scenario 6, Run 4 (green) and Run 5 (orange).

3.6.3 Overall Results

The iterative process and run-to-run comparisons described in the previous section resulted in five runs that demonstrate the interpretation of the stability of simulation results with respect to sampling options.

An ideal selection of sampling options leads to modeling results that converge within an acceptable computational time. Figure 122 shows the results for the mean probability of occurrence of circumferential rupture for Runs 1 through 5. Though there appears to be a trend of higher approximations with SRS, this is an artifact of the random seed choice as studied in-depth in Section 4.4. For this scenario, the other outputs of interest converge comparably to those for the occurrence of circumferential ruptures.

All of the other sampling schemes result in approximations that are nearly equal. This may suggest that the benefits provided by using LHS and/or IS lead to convergence more quickly than SRS alone. If this is the case, estimates obtained using only SRS may eventually converge to the other results if the sample size is large enough. This determination, however, requires that the sample size necessary for this convergence is computationally feasible. The difference between the SRS

results and the other results does decrease for all outputs as the epistemic sample size increases from 100 to 500, as seen in Figure 122. If this pattern continues, the SRS results will eventually converge to the other results.

The uncertainty associated with the result of interest, shown in the form of 95% confidence intervals, can also be used to identify beneficial sampling selections for this scenario. Figure 114 and Figure 116 show how uncertainty is affected by the epistemic sample size. Choosing the final sample size will depend on both the computational time available and the level of uncertainty that the user is willing to accept. The uncertainty continuously decreases with increasing sample size, but the incremental gain from 500 to 1000 may not be worthwhile to a user who is concerned about computational time. Depending on user-needs pertaining to mean precision, smaller epistemic sample sizes may be sufficient.

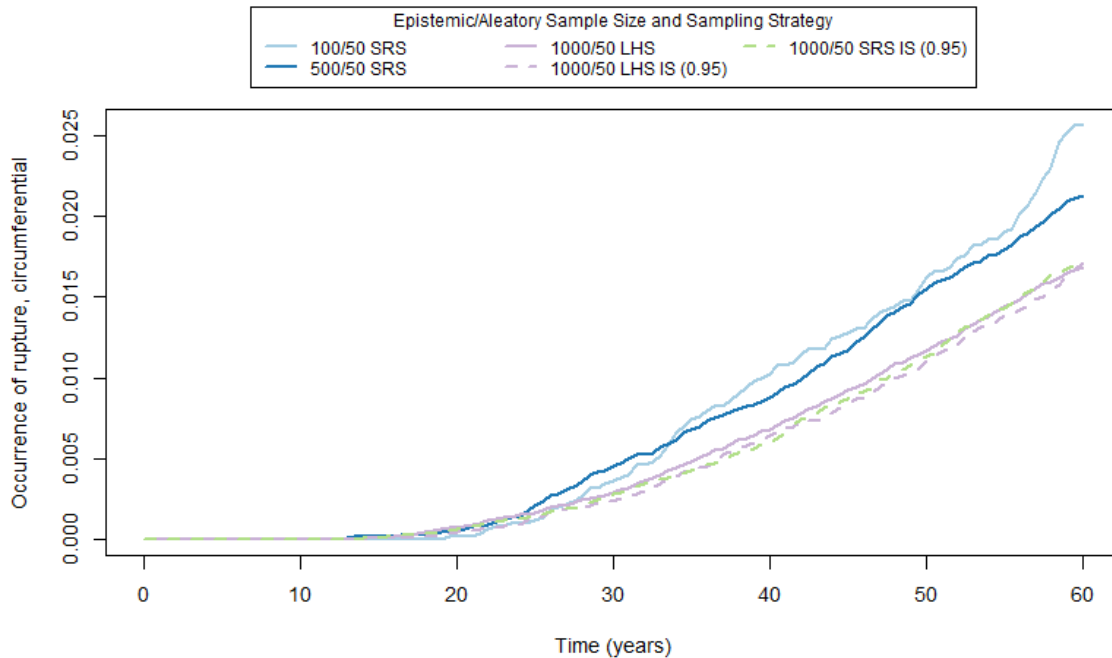


Figure 122: Mean probability of occurrence of circumferential rupture for Scenario 6, Runs 1 through 5.

Importance sampling had more of an impact on decreasing the uncertainty in the result of interest for runs using SRS and runs with smaller epistemic sample sizes than for results using LHS and large sample sizes. Importance sampling does not necessarily impact the uncertainty in the result of interest using these sampling options; the impact of importance sampling on uncertainty is highly dependent upon the quantile chosen for sampling and the magnitude of the quantity being estimated. As in Scenario 3, the lack of gain in precision using importance sampling is likely due to the relatively high frequency of event occurrence for this scenario; the benefit from applying importance sampling will increase as the probability of event occurrence decreases. Importance sampling can be used with no additional computational cost and allows for a broader coverage of the output space, particularly if estimation of quantiles of the epistemic CDFs is of interest. However, importance sampling does increase the complexity of the data analysis, as all inferences

on the data must take into account the importance weight for each epistemic sample. Further, importance sampling can decrease efficiency in QoI estimates with poor selection of the importance distribution. These considerations should be weighed when deciding if importance sampling should be used for a specific analysis.

In this analysis, the sampling scheme for Run 4 (1000 epistemic, 50 aleatory, LHS and IS) provided reasonable results. Because this scheme also exercises many of the sampling options analysts may elect to use, it is further studied in the following sections.

3.6.4 Epistemic Uncertainty Analysis

An uncertainty analysis was conducted for Run 4 with 1000 epistemic and 50 aleatory samples with LHS and IS for Scenario 6. Plots of the estimated probabilities of occurrence of circumferential crack, leak, and rupture are displayed in Figure 123, Figure 124, and Figure 125, respectively. These figures show the probability of occurrence for each epistemic realization (grey lines), the mean (red line), and the 5th, 50th, and 95th percentiles (blue, green, and purple lines, respectively). In contrast to the 95% confidence intervals for the mean shown in Section 3.6.2, which were calculated using the bootstrap method, the percentiles shown here offer a summary of the total range of results. They encompass 90% of all the epistemic realizations observed in this run. For example, the 50th percentile for circumferential crack is about 0.02 at year 60, meaning that half of the epistemic realizations have about a 2% chance of circumferential crack by year 60. For leak and rupture, the 5th and 50th percentiles are uniformly zero up to year 60, indicating that at least 50% of the realizations did not experience a leak or rupture in the first 60 years. These results are consistent with the results from Scenario 3, showing a slightly decreased probability of circumferential rupture or leak by year 60 with Zn mitigation at 20 years.

As seen in the Scenario 3 results, the epistemic realizations for leak (Figure 124) are also slightly shifted to the right in Scenario 6 compared to crack (Figure 123) and ruptures (Figure 125) are slightly shifted to the right compared to cracks. This intuitive result indicates that for this scenario, cracks are soon followed by leaks and leaks are soon followed by ruptures. This relationship is further illustrated in Figure 126, which displays the mean and 95th percentile for leak and crack, and in Figure 127, which displays the mean and 95th percentile for leak and rupture. These plots also show that the spread between the mean and the 95th percentile is larger for crack than it is for leak and larger for leak than for crack – indicating rupture results are relatively more certain than leak results and leak results are relatively more certain than crack results, though this difference may not be significant.

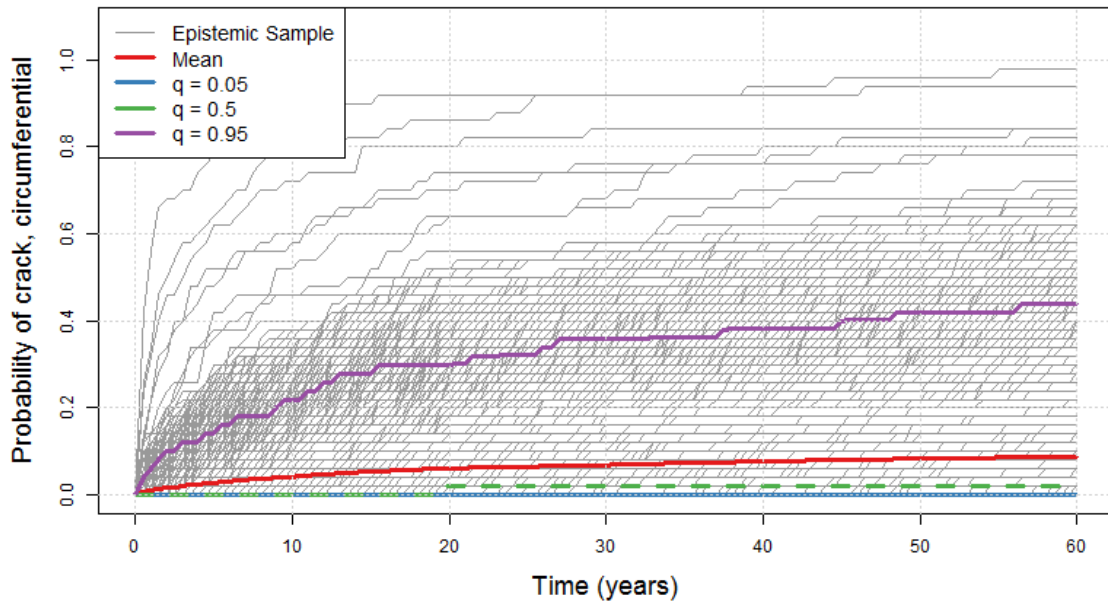


Figure 123: Probability of occurrence of circumferential crack for each epistemic realization (grey), the mean (red), and the 5th (blue), 50th (green), and 95th (purple) percentiles for Scenario 6, Run 4.

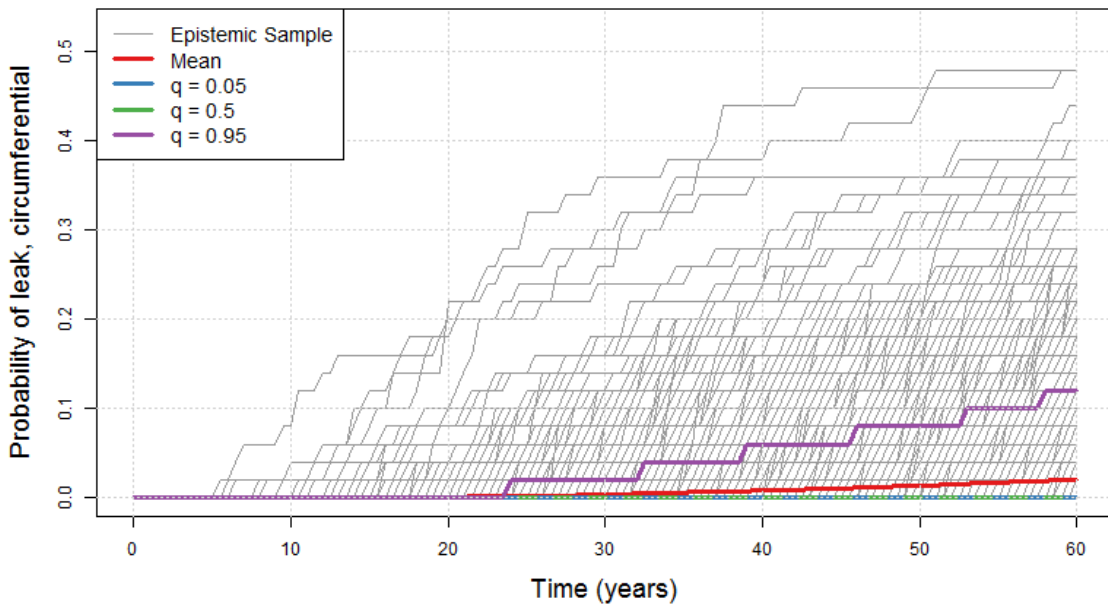


Figure 124: Probability of occurrence of circumferential leak for each epistemic realization (grey), the mean (red), and the 5th (blue), 50th (green), and 95th (purple) percentiles for Scenario 6, Run 4.

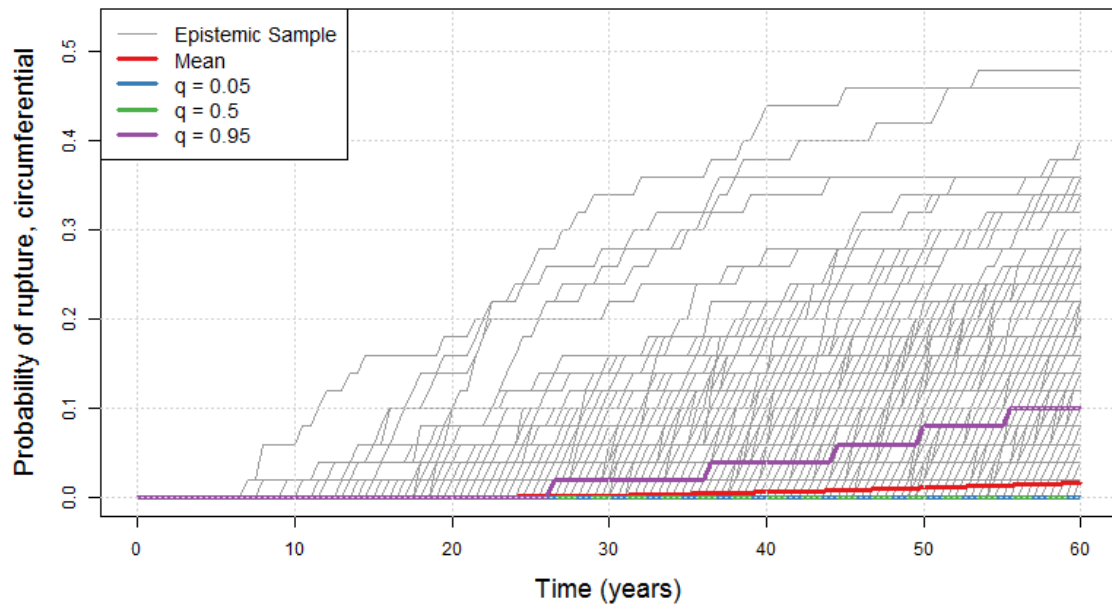


Figure 125: Probability of occurrence of circumferential rupture for each epistemic realization (grey), the mean (red), and the 5th (blue), 50th (green), and 95th (purple) percentiles for Scenario 6, Run 4.

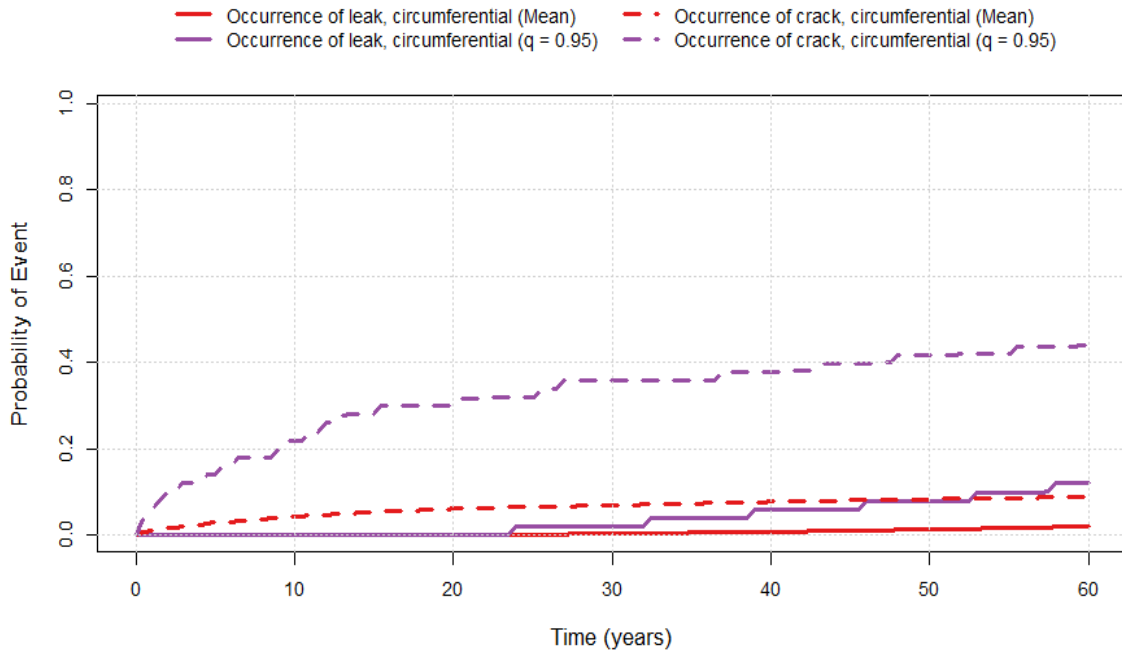


Figure 126: Mean (red) and 95th percentile (purple) for occurrence of circumferential leak (solid) and crack (dashed) for Scenario 6, Run 4.

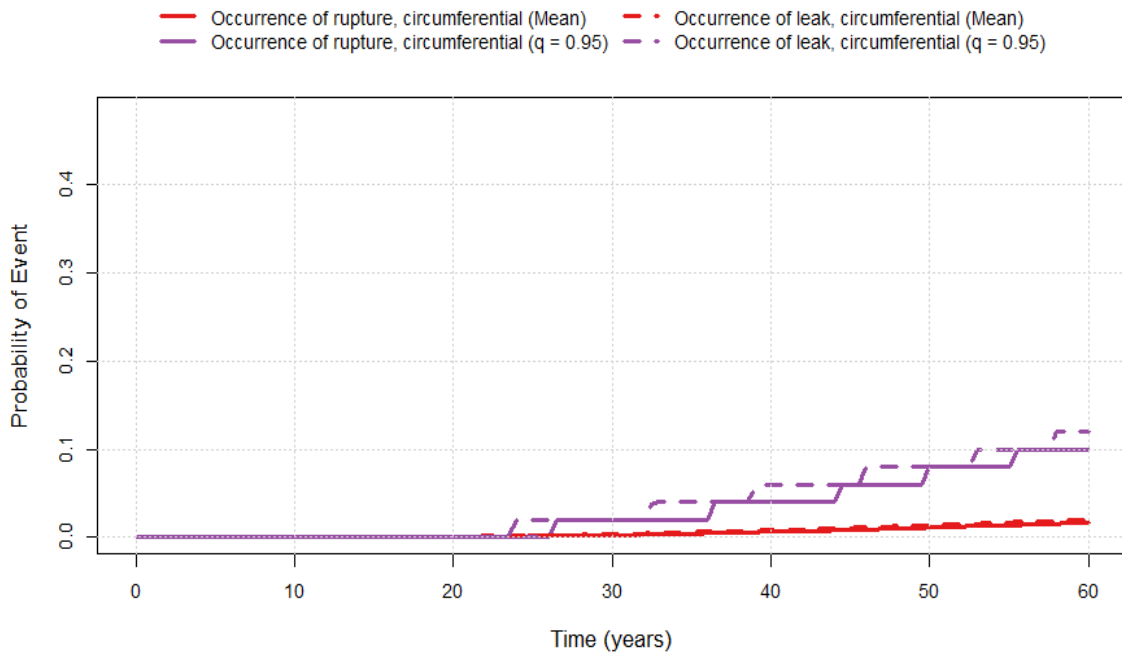


Figure 127: Mean (red) and 95th percentile (purple) for occurrence of circumferential rupture (solid) and leak (dashed) for Scenario 6, Run 4.

3.6.5 Convergence Analysis

In order to gain better insight into the convergence of the sampling options for Run 4 (1000 epistemic samples with LHS and IS), this sampling scheme was re-run five times with different random number seeds for both the epistemic and aleatory loops. Each replicate was used to give an estimate of the mean probability of occurrence of circumferential ruptures. Figure 128 shows a comparison of these estimates. The replicates produce similar estimates within a range with width less than 0.007 at 60 years. Over time, the fifth replicate experienced a higher occurrence of circumferential ruptures than the other four replicates.

The five replicates were also used to estimate an overall mean value and a 95% prediction interval was created around this mean, as seen in Figure 129. The interval widens over the course of the simulation time, though the final width is still around 0.01. If that level of uncertainty is acceptable for the analysis, the sampling scheme converges sufficiently. If that level of uncertainty is not acceptable, the sample size should be increased or the importance sampling quantile should be further refined.

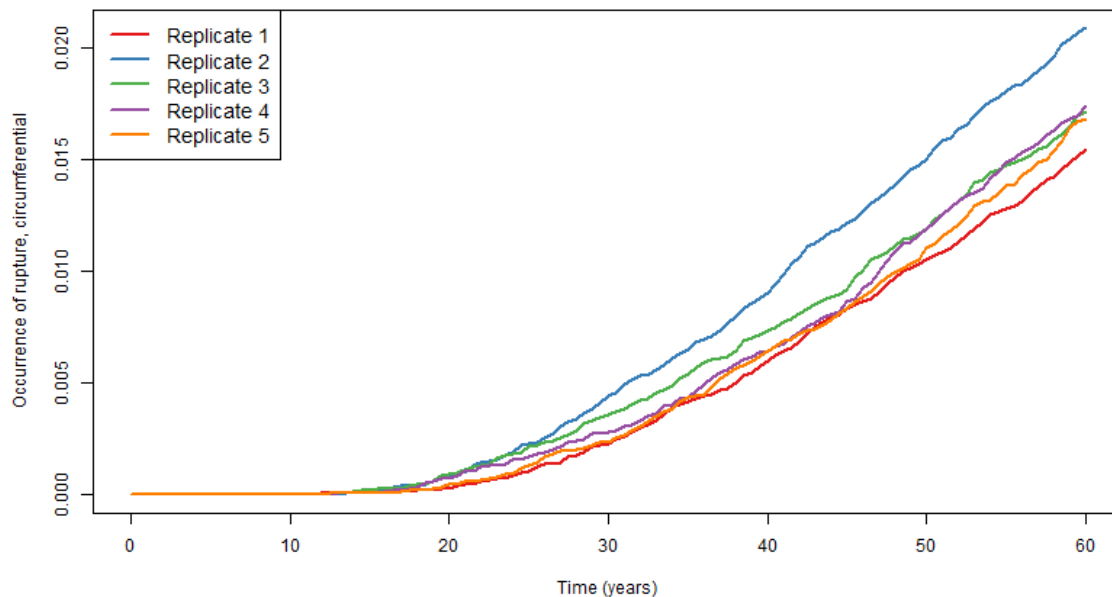


Figure 128: Comparison of estimated mean probability of occurrence of circumferential ruptures for five convergence replicates for Scenario 6, Run 4.

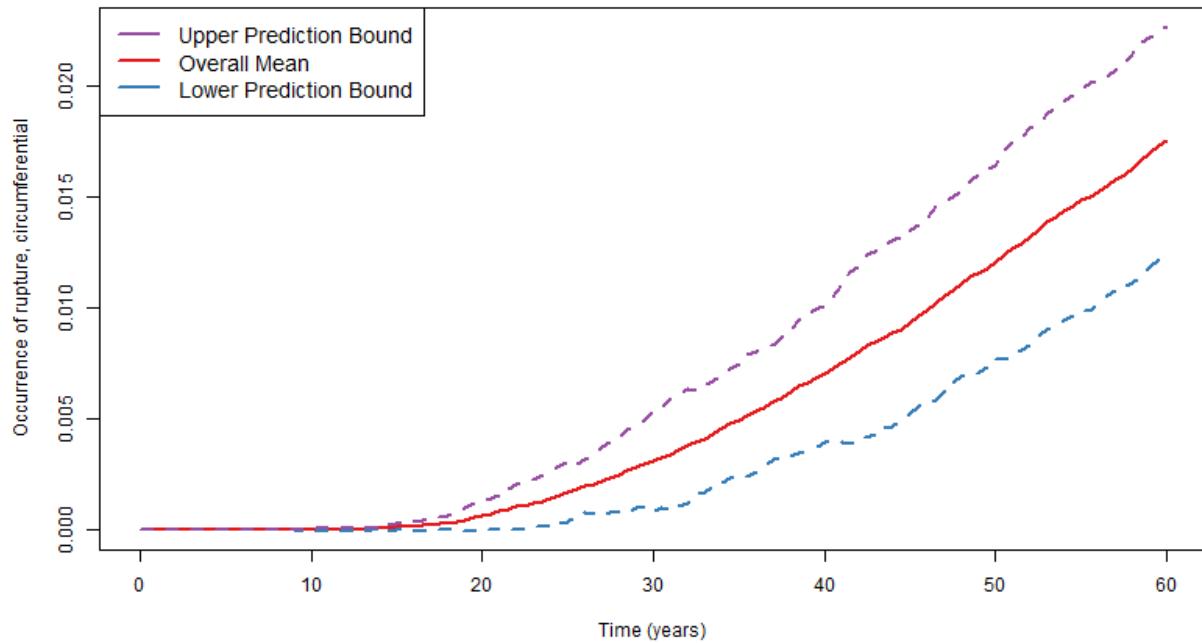


Figure 129: 95% prediction interval over mean probability of occurrence of circumferential ruptures for five replicates for Scenario 6, Run 4.

3.6.6 Comparison to Scenario 3 Results

The Scenario 6 results were compared to the Scenario 3 results to examine the effect of Zn mitigation at 20 years. The runs compared (Scenario 3 Run 9 and Scenario 6 Run 4) have the same sampling options (LHS IS 1000/50), the same epistemic and aleatory random seeds, and identical inputs except for the inclusion of Zn mitigation at 20 years in Scenario 3. The purpose of Zn mitigation is to mitigate crack initiation. This effect is illustrated in Figure 130, which compares the mean occurrences of axial cracks (dashed lines) and circumferential cracks (solid lines) between the two scenarios. The probability of occurrence of axial cracks by year 60 is reduced by about 0.025 and the probability of occurrence of circumferential cracks is reduced by about 0.015.

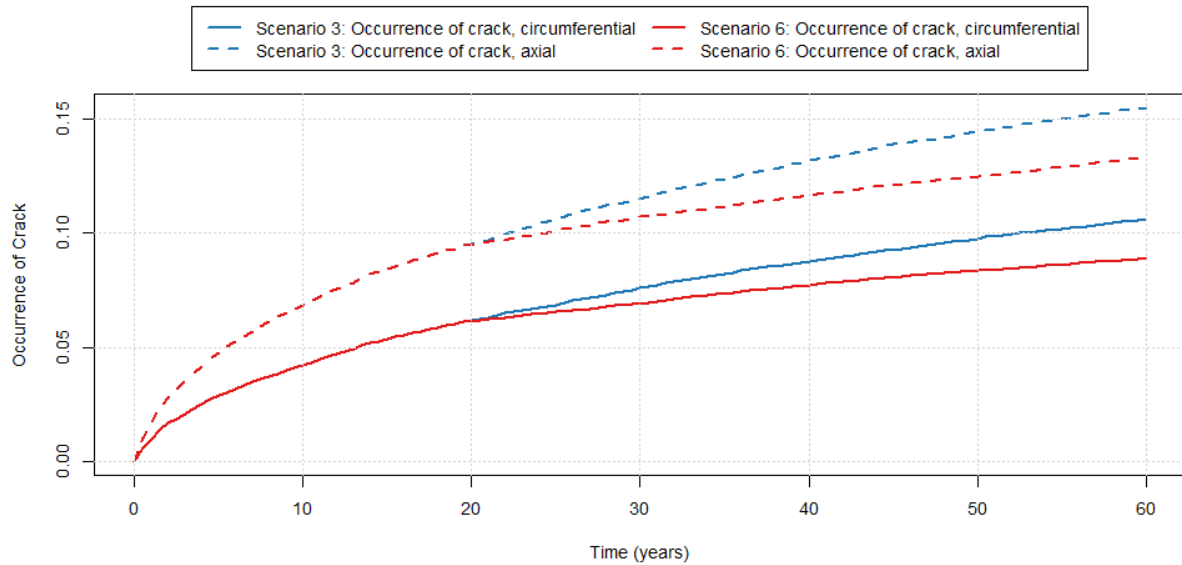


Figure 130: Mean probability of axial (dashed lines) and circumferential (solid lines) crack occurrence for Scenario 3 (blue lines) and Scenario 6 (red lines).

The inclusion of Zn mitigation can be seen in the divergence of the red and blue curves at approximately 20 years. Prior to mitigation, Scenario 3 and Scenario 6 estimate the same probabilities of occurrence of cracks. Beyond 20 years, cracks still occur in both scenarios, but the rate of occurrence is decreased in Scenario 6.

Though Zn mitigation affects crack formation, not leaks, the estimates of total leak rate from Scenario 3 and Scenario 6 appear to similarly deviate, though by a smaller magnitude, in Figure 131. This effect is a result of a lower probability of cracks; fewer cracks leads to fewer leaks. Similarly, the probability of circumferential ruptures is also slightly lower in Scenario 6, shown in Figure 132.

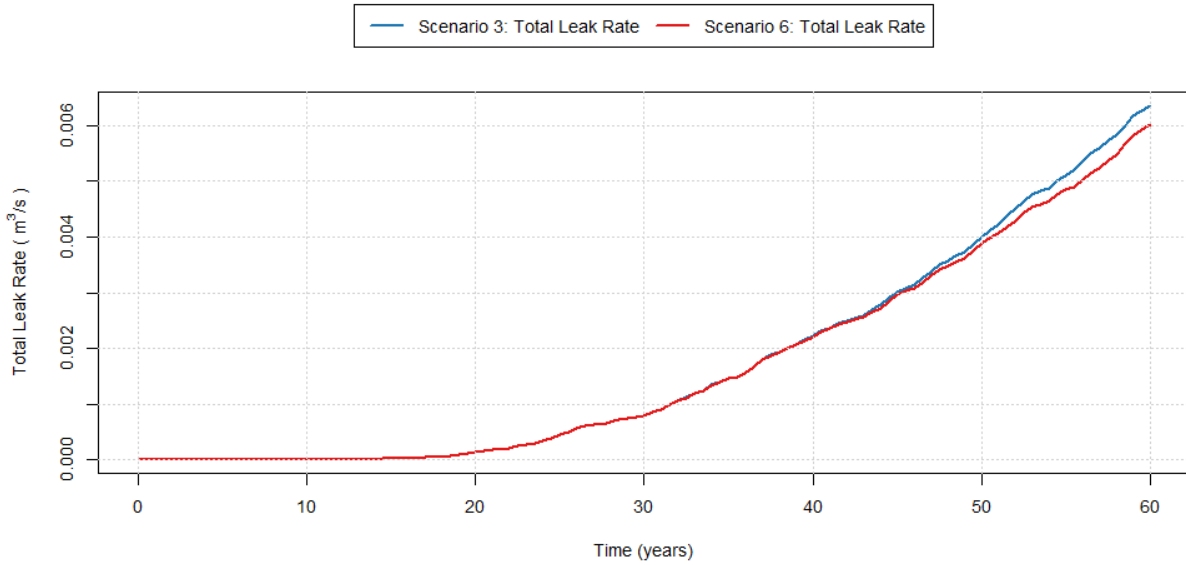


Figure 131: Mean estimate of total leak rate for Scenario 3 (blue) and Scenario 6 (red).

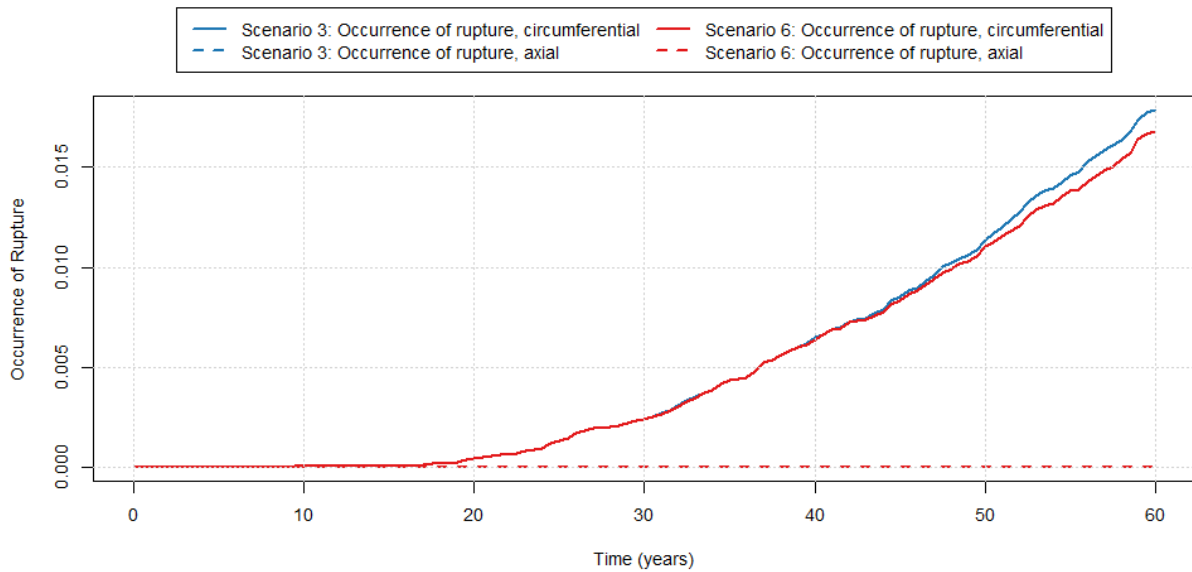


Figure 132: Mean estimate of the probability of occurrence of rupture for Scenario 3 (blue) and Scenario 6 (red).

The results for the other quantities of interest are similar, showing small reductions in crack dimensions, leaks, and ruptures as a result of cracks forming later in the simulation having less time to grow.

3.7 Scenario 7

3.7.1 Scenario 7 Summary

The results and conclusions of a detailed analysis of sampling schemes and their impact on an understanding of the uncertainty found in the simulation results of Scenario 7 are given in this section.

Scenario 7 uses the following defining options in the xLPR code:

- PWSCC flaw initiation
- Circumferential and axial flaw orientation
- PWSCC flaw growth
- Chemical mitigation with dissolved hydrogen (H_2) at 20 years

This scenario is based on Scenario 3 with the addition of H_2 chemical mitigation.

The outputs of interest defined for this scenario for both circumferential and axial cracks are:

- Probability of occurrence of crack
- Probability of occurrence of leak
- Probability of occurrence of rupture
- Total Leak Rate
- Maximum crack depth
- Maximum crack inner half-length
- Maximum crack outer half-length

The comprehensive list of simulations that were used in the analysis of this scenario is shown below. A constant aleatory sample size of 50 was used for each of the sampling schemes considered.

1. 100/50 (epistemic/aleatory sample size) with simple random sampling (SRS)
2. 500/50 with simple random sampling (SRS)
3. 1000/50 with Latin hypercube sampling (LHS)
4. 1000/50 with LHS and importance sampling (IS) on p2543 with target quantile of 0.95
5. 1000/50 with SRS and IS on p2543 with target quantile of 0.95

Sampling schemes were chosen based on insight gained from the results of the Scenario 3 analysis and to exercise all of the sampling options for epistemic uncertainty.

The results of the Scenario 7 analysis are consistent with the analysis of Scenario 3. The only difference between the scenarios is that Scenario 7 includes H_2 mitigation at 20 years; in every other respect the scenario is identical to Scenario 3. While many outputs of interest were defined for this scenario, the results of regression analyses indicate that the same variable, the multiplier on Direct Model 1 (DM1) proportionality constant A for the PWSCC crack initiation module, hereafter referred to as p2543, was the most important input for explaining variability in all of the outputs for Scenario 7. For the sake of brevity, the only result shown in this section is the occurrence of circumferential ruptures, as the results for other outputs are comparable with respect to convergence and stability. This output also demonstrates the effect of H_2 mitigation as it is expected to decrease

the probability of ruptures (and leaks and decrease crack dimensions and leak rates) by delaying crack growth. When used, importance sampling was applied to p2543 to improve the precision in the output of interest.

As in the Scenario 3 analysis, overall, results confirm that increasing the epistemic sample size decreases the uncertainty in the mean of the output (e.g., probability of occurrence of ruptures from circumferential cracks) as expected, especially when moving from 100 to 500 epistemic samples. The decrease in uncertainty in the mean of the output is also evident when transitioning from 500 to 1000 epistemic realizations. An epistemic sample size of 1000 may be sufficient for this scenario, though a sample of 500 may also provide sufficient mean precision depending on user needs.

3.7.2 Analysis Progression

The five sampling schemes, informed by the Scenario 3 analysis described in Section 3.3, were chosen to demonstrate stability and the effect of sampling options of the stability of results. These schemes progress in sample size and sampling options.

Runs 1 and 2 are identical in all sampling options except the epistemic sample size. The epistemic sample size for Run 1 was 100 and this was increased to 500 for Run 2. The mean probability of occurrence of circumferential ruptures from Runs 1 and 2 is plotted in Figure 133 with 95% confidence intervals calculated using bootstrapping. The mean estimates for both runs are fairly close initially, but deviate at about 15 years and continue to differ through the remainder of the simulation. This suggests that larger sample sizes are needed to better approximate the reduction in occurrence of ruptures when H₂ mitigation is used.

With regard to stability, the width of the confidence intervals around the estimated mean decreases when the epistemic sample size is increased to 500. This is shown in Figure 134. The width of the confidence interval decreases by about 0.008 when the epistemic sample size is increased from 100 to 500. This suggests that more samples are necessary if greater precision is required around the mean estimate.

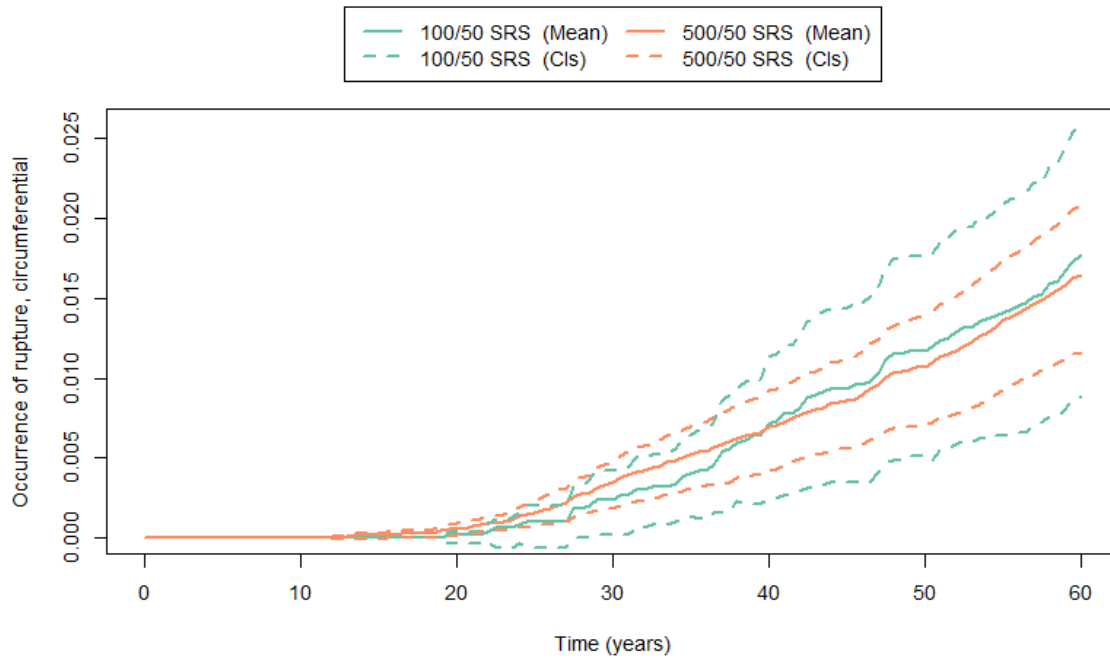


Figure 133: Mean probability of occurrence of circumferential ruptures (solid lines) and 95% confidence intervals (dashed lines) for Scenario 7, Run 1 (green) and Run 2 (orange).

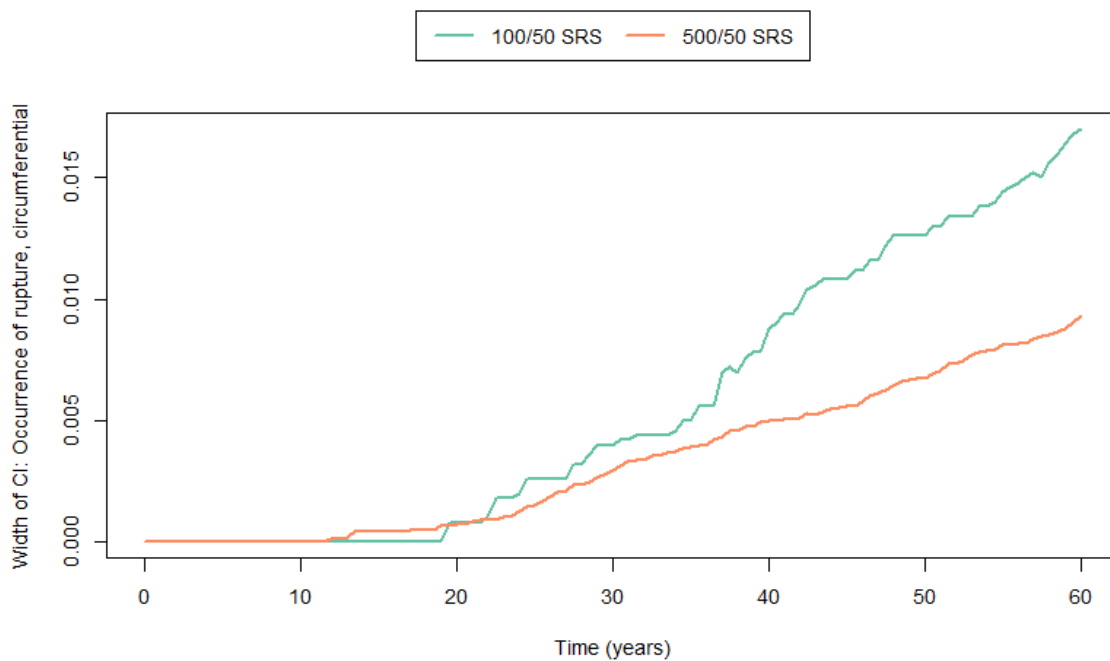


Figure 134: Width of 95% confidence interval for the estimate of mean probability of occurrence of circumferential ruptures for Scenario 7, Run 1 (green) and Run 2 (orange).

The Scenario 3 analysis showed that LHS sampling could improve stability and there is no computational reason to prefer SRS to LHS. Thus, to improve upon the results in Run 2, the epistemic sample size was increased to 1000 and LHS was used instead of SRS to sample epistemically uncertain parameters for Scenario 7.

The increase in sample size and change in sampling strategy from Run 2 to Run 3 further decreased the width of the confidence interval by about 0.004, shown in Figure 135 and Figure 136. This supports both the use of a larger sample size and LHS. The divergence of the estimates from Run 2 to Run 3 motivates the use of importance sampling.

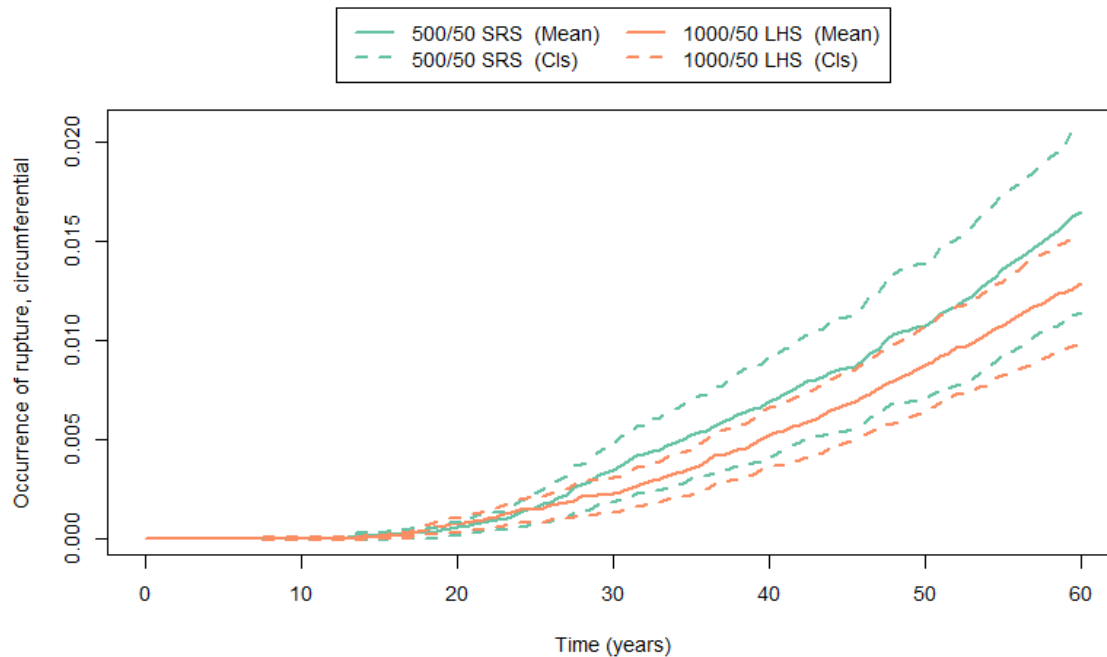


Figure 135: Mean probability of occurrence of circumferential ruptures (solid lines) and 95% confidence intervals (dashed lines) for Scenario 7, Run 2 (green) and Run 3 (orange).

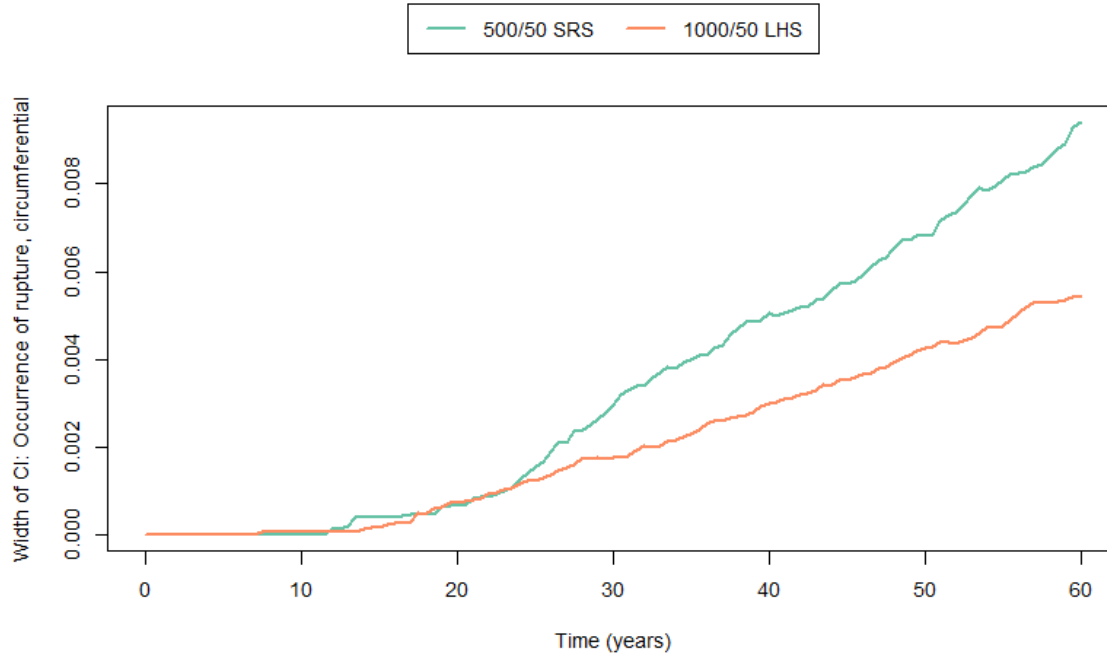


Figure 136: Width of the 95% confidence interval for probability of occurrence of circumferential ruptures for Scenario 7, Run 2 (green) and Run 3 (orange).

Recall that importance sampling concentrates sampling more in regions of the input space that are expected to contribute most to variation in the simulation results. Stepwise rank regression results from Run 3 on the occurrence of circumferential rupture and other outputs of interest for Scenario 7 are presented in Table 8. These results are consistent with the regression results from Runs 1 and 2. The parameter identified in Runs 1, 2, and 3 for Scenario 7 as the most important variable with respect to variability in the probability of occurrence of circumferential ruptures is p2543, the multiplier on DM1 proportionality constant A. The influence of this variable can be seen both by the high SRRC value (0.519) and the scatter plots of p2543 with the occurrence of circumferential rupture in Figure 137. The increasing occurrence of circumferential ruptures as p2543 increases is consistent in all three runs. Hence, p2543 is a good candidate for IS for Scenario 7. The result stratification shown in Figure 137 is a result of the selected sample size. These results represent an average over a fixed number of indicator functions (value of 0 or 1). Thus, the probability can only occur incrementally as a function of the aleatory sample size.

Table 8: Summary of the stepwise rank regression analysis results for all outputs at year 60 for Scenario 7, Run 3.

Variable Identifier	Variable Name	SRRC	Variable Identifier	Variable Name	SRRC
Probability of Occurrence of Axial Cracks – $R^2 = 0.886$			Maximum Axial Crack Depth – $R^2 = 0.818$		
p2543	Multiplier proport. Const. A (DM1)	0.926	p2543	Multiplier proport. Const. A (DM1)	0.856
p4350	Hoop WRS Pre-mitigation	-0.182	p4350	Hoop WRS Pre-mitigation	-0.235
p1102	Pipe Wall Thickness	-0.053	p2592	Comp-to-Comp Variab Factor, fcomp	0.103
p2591	Activation Energy, Qg	0.035	p2595	Charact Width of Peak vs ECP, c	0.101
p9002	Surface Crack Dist Rule Modifier	-0.029	p1102	Pipe Wall Thickness	-0.089
Probability of Occurrence of Circumferential Cracks – $R^2 = 0.843$			Maximum Circumferential Crack Depth – $R^2 = 0.807$		
p2543	Multiplier proport. Const. A (DM1)	0.912	p2543	Multiplier proport. Const. A (DM1)	0.873
p4352	Axial WRS Pre-mitigation	-0.110	p4352	Axial WRS Pre-mitigation	-0.153
p5105	a (circ)	0.049	p1102	Pipe Wall Thickness	-0.074
p5106	b (circ)	0.043	p2595	Charact Width of Peak vs ECP, c	0.074
p1102	Pipe Wall Thickness	-0.034	p2594	Peak-to-Valley ECP Ratio - 1, P-1	-0.068
Probability of Axial Leak – $R^2 = 0.608$			Maximum Circumferential Inner Half-Length – $R^2 = 0.788$		
p2543	Multiplier proport. Const. A (DM1)	0.693	p2543	Multiplier proport. Const. A (DM1)	0.858
p4350	Hoop WRS Pre-mitigation	-0.202	p4352	Axial WRS Pre-mitigation	-0.158
p2592	Comp-to-Comp Variab Factor, fcomp	0.200	p2592	Comp-to-Comp Variab Factor, fcomp	0.090
p2595	Charact Width of Peak vs ECP, c	0.196	p2595	Charact Width of Peak vs ECP, c	0.090
p2594	Peak-to-Valley ECP Ratio - 1, P-1	-0.155	p2594	Peak-to-Valley ECP Ratio - 1, P-1	-0.084
Probability of Circumferential Leak – $R^2 = 0.449$			Maximum Axial Inner Half-Length – $R^2 = 0.875$		
p2543	Multiplier proport. Const. A (DM1)	0.566	p2543	Multiplier proport. Const. A (DM1)	0.912
p2592	Comp-to-Comp Variab Factor, fcomp	0.192	p4350	Hoop WRS Pre-mitigation	-0.215
p4352	Axial WRS Pre-mitigation	-0.185	p1102	Pipe Wall Thickness	-0.059
p1102	Pipe Wall Thickness	-0.153	p2591	Activation Energy, Qg	0.041
p2594	Peak-to-Valley ECP Ratio - 1, P-1	-0.149	p9002	Surface Crack Dist Rule Modifier	-0.028
Probability of Circumferential Rupture – $R^2 = 0.410$			Maximum Axial Outer Half-Length – $R^2 = 0.608$		
p2543	Multiplier proport. Const. A (DM1)	0.519	p2543	Multiplier proport. Const. A (DM1)	0.692
p2592	Comp-to-Comp Variab Factor, fcomp	0.209	p4350	Hoop WRS Pre-mitigation	-0.202
p1102	Pipe Wall Thickness	-0.179	p2592	Comp-to-Comp Variab Factor, fcomp	0.200
p2594	Peak-to-Valley ECP Ratio - 1, P-1	-0.173	p2595	Charact Width of Peak vs ECP, c	0.195
p4352	Axial WRS Pre-mitigation	-0.166	p2594	Peak-to-Valley ECP Ratio - 1, P-1	-0.154
Total Leak Rate – $R^2 = 0.605$			Maximum Circumferential Outer Half-Length – $R^2 =$		
p2543	Multiplier proport. Const. A (DM1)	0.695	p2543	Multiplier proport. Const. A (DM1)	0.564
p2592	Comp-to-Comp Variab Factor, fcomp	0.214	p2592	Comp-to-Comp Variab Factor, fcomp	0.197
p2595	Charact Width of Peak vs ECP, c	0.183	p4352	Axial WRS Pre-mitigation	-0.184
p2594	Peak-to-Valley ECP Ratio - 1, P-1	-0.159	p2594	Peak-to-Valley ECP Ratio - 1, P-1	-0.155
p1102	Pipe Wall Thickness	-0.124	p1102	Pipe Wall Thickness	-0.149

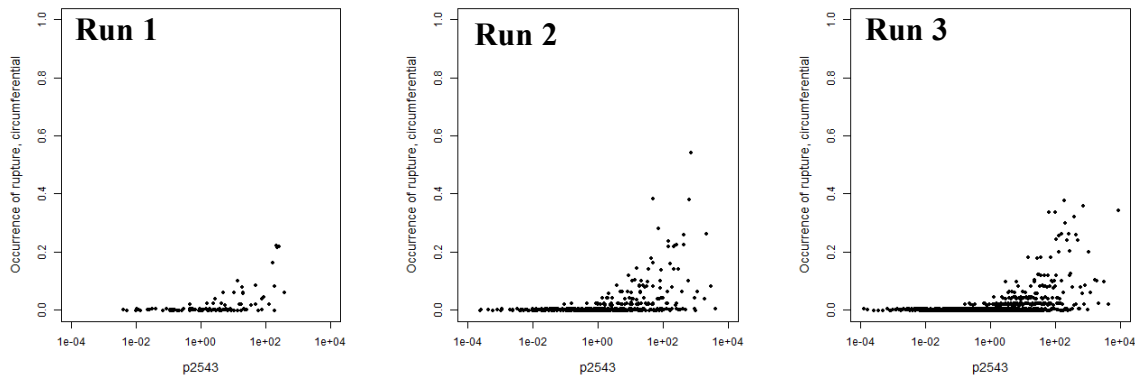


Figure 137: Scatter plots for variable p2543 (multiplier on DM1 proportionality constant A) and the probability of occurrence for circumferential ruptures for Scenario 7, Run 1 (left), Run 2 (center), and Run 3 (right).

Run 4 improves upon Run 3 by incorporating importance sampling, while maintaining the same sample size (1000) and sampling strategy (LHS). Importance sampling was applied on p2543 at the

95th quantile to oversample the upper tail of the p2543 distribution to increase the occurrence of circumferential ruptures. The comparison between mean estimates of the probability of occurrence of circumferential ruptures for Runs 3 and 4 is plotted in Figure 138, showing a small effect from importance sampling early in the simulation, but no reduction in uncertainty from IS later in the simulation. The use of importance sampling for this scenario may not be necessary when the output of interest is the occurrence of circumferential ruptures because importance sampling is most beneficial for reducing uncertainty around estimates of less likely events.

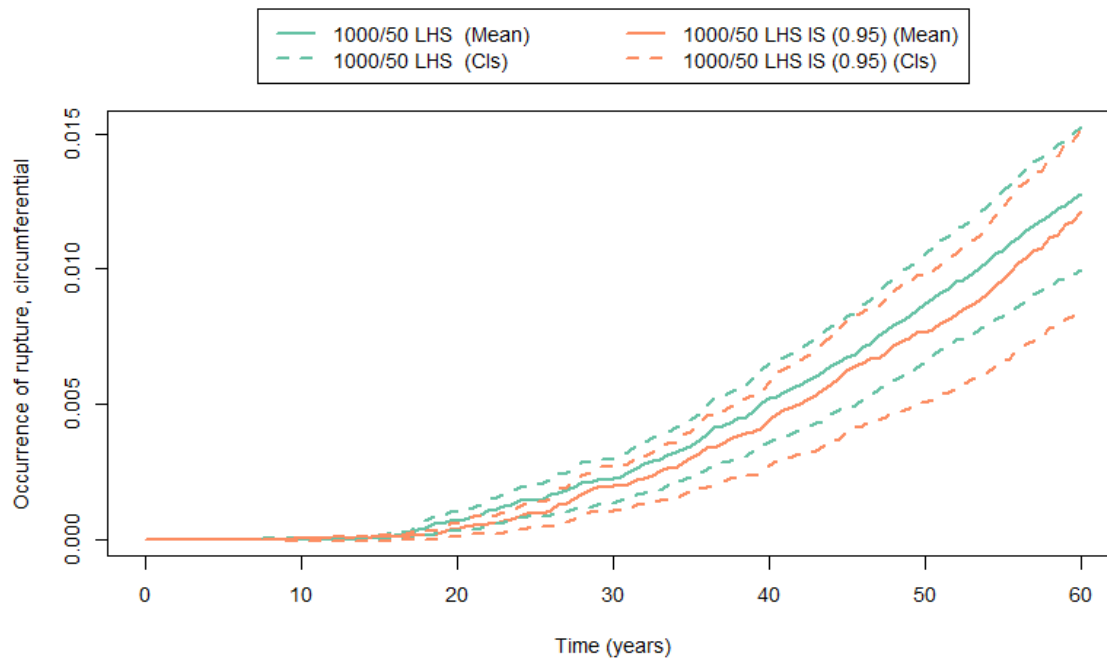


Figure 138: Mean probability of occurrence of circumferential ruptures (solid line) and 95% confidence intervals (dashed lines) for Scenario 7, Run 3 (green) and Run 4 (orange).

Stability can be further investigated using Run 5, which maintains the sample size and importance sampling on p2543 from Run 4, but employs SRS rather than LHS. With SRS, the results are comparable to the comparison between Run 3 and Run 4. The estimated mean occurrence of circumferential ruptures with confidence intervals from Run 5 are plotted with those from Run 4 in Figure 139 to show these differences.

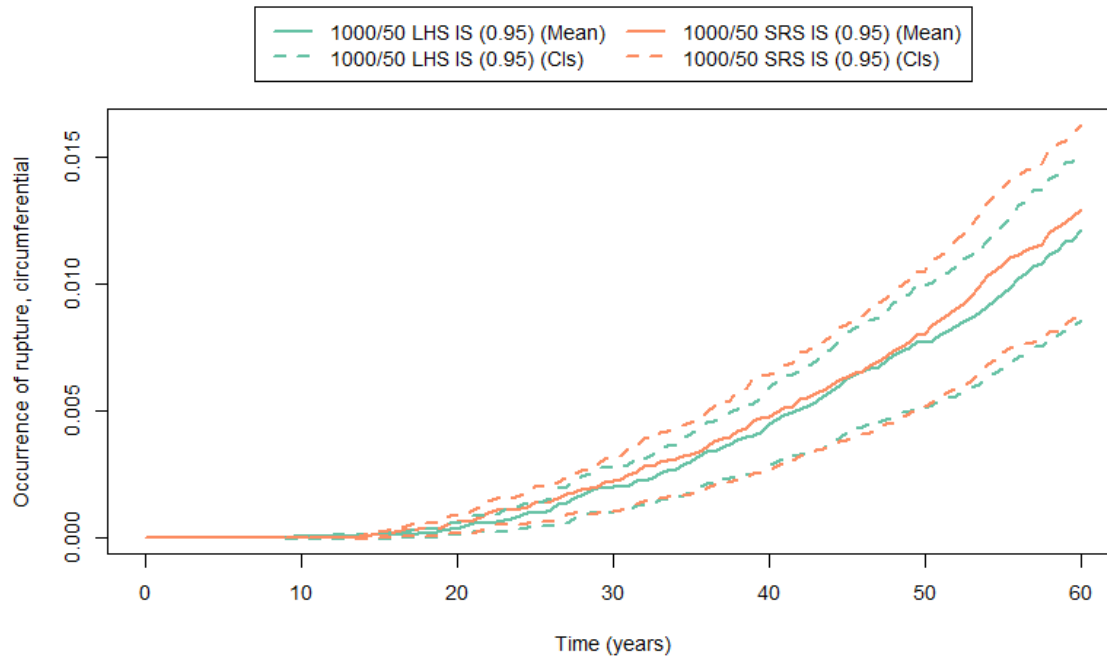


Figure 139: Mean Probability of occurrence of circumferential ruptures (solid line) and 95% confidence intervals (dashed lines) for Scenario 7, Run 4 (green) and Run 5 (orange).

3.7.3 Overall Results

The iterative process and run-to-run comparisons described in the previous section resulted in five runs that characterize the stability of simulation results with respect to sampling options. A comparison across this collection of runs can be used to provide guidance on how to choose a sampling scheme for Scenario 7.

An ideal selection of sampling options leads to modeling results that converge within an acceptable computational time. Figure 140 shows the results for the mean probability of occurrence of circumferential ruptures for Runs 1 through 5. The results show a large reduction in variation between estimates as the epistemic sample size is increased beyond 500 and importance sampling is used. If the range of results is sufficient for applications, this may be acceptable. However, if the range of variation is too large, then the epistemic sample size should be increased further.

The uncertainty associated with the result of interest, shown in the form of 95% confidence intervals, can also be used to identify beneficial sampling selections for this scenario. Figure 133 and Figure 135 show how uncertainty is affected by the epistemic sample size. Choosing the final sample size will depend on both the computational time available and the level of uncertainty that the user is willing to accept. The uncertainty continuously decreases with increasing sample size and sample sizes above 1000 may be worthwhile for estimating the probability of circumferential ruptures in this scenario if high precision is needed.

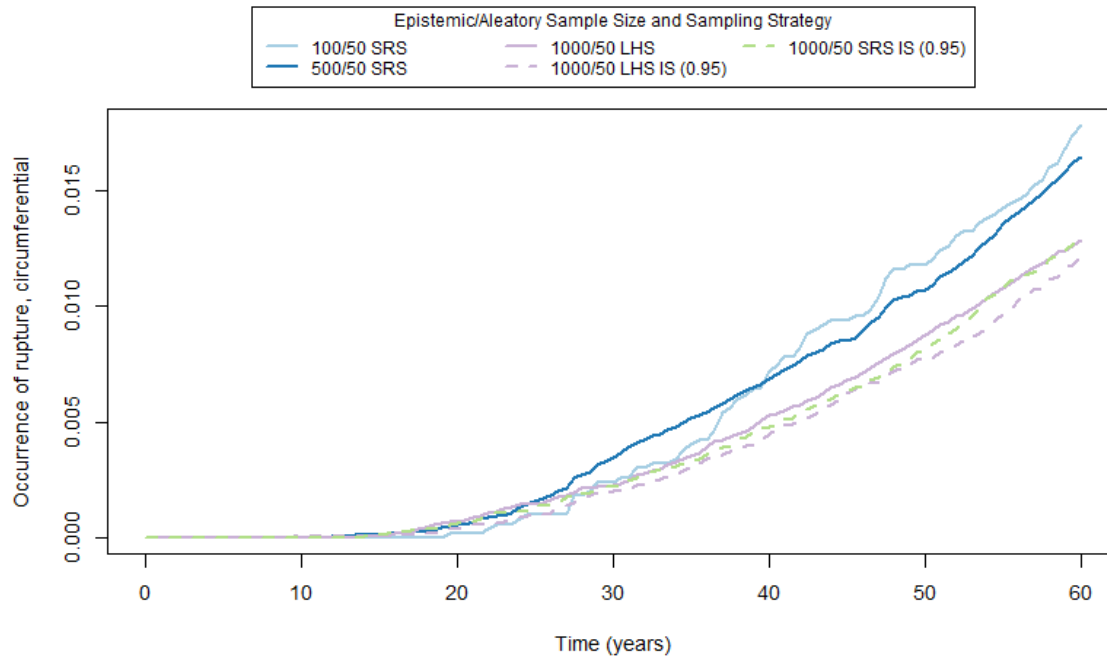


Figure 140: Mean probability of occurrence circumferential ruptures for Scenario 7, Runs 1 through 5.

Importance sampling does not decrease the uncertainty in the result of interest for runs using either SRS or LHS. This may be a result of the relatively high probability being approximated or an indication that the quantile chosen for importance sampling is not sufficient for reducing uncertainty.

In this analysis, the sampling scheme for Run 4 (1000 epistemic, 50 aleatory, LHS and IS) provided reasonable confidence interval results. Because this scheme also exercises many of the sampling options analysts may elect to use, it is further studied in the following sections.

3.7.4 Epistemic Uncertainty Analysis

An uncertainty analysis was conducted for Run 4 with 1000 epistemic and 50 aleatory samples with LHS and IS for Scenario 7.

Plots of the estimated probabilities of occurrence of circumferential crack, leak and rupture are displayed in Figure 141, Figure 142, and Figure 143, respectively. These figures show the probability of occurrence for each epistemic realization (grey lines), the mean (red line), and the 5th, 50th, and 95th percentiles (blue, green, and purple lines, respectively). In contrast to the 95% confidence intervals for the mean shown in Section 3.7.2, which were calculated using the bootstrap method, the percentiles shown here offer a summary of the total range of results. They encompass 90% of all the epistemic realizations observed in this run. For example, the 50th percentile for circumferential crack is about 0.02 at year 60, meaning that half of the epistemic realizations have about a 2% chance of circumferential crack by year 60. For leak and rupture, the 5th and 50th percentiles are also uniformly zero up to year 60, indicating that at least 50% of the realizations are not expected to experience a leak or rupture in the first 60 years. These results are consistent with

the results from Scenario 3, showing a slightly decreased probability of circumferential leaks and ruptures by year 60 with H₂ mitigation at 20 years.

As seen in the Scenario 3 results, the epistemic realizations for leaks are also slightly shifted to the right in Scenario 7 compared to cracks and ruptures are shifted to the right compared to leaks. This intuitive result indicates that for this scenario, cracks are eventually followed by leaks and leaks are eventually followed by ruptures; H₂ mitigation may delay leaks, but when cracks occur, leaks and ruptures follow. This relationship is further illustrated in Figure 144 and Figure 145, which display the mean and 95th percentile for crack and leak and for leak and rupture, respectively. Figure 144 shows that the spread between the mean and the 95th percentile is larger for crack (dashed lines) than it is for leaks (solid lines) and Figure 145 shows that the spread is larger for leaks than for ruptures— indicating crack results are relatively more uncertain than leak results and leak results are relatively more uncertain than rupture results, as in Scenario 3. The mean probability of occurrence for crack, rupture and leak are very similar in Scenario 3 and Scenario 7, though the mean for leaks and ruptures for Scenario 7 are shifted down as expected. The mean for occurrence of crack is unchanged as H₂ mitigation does not affect crack initiation.

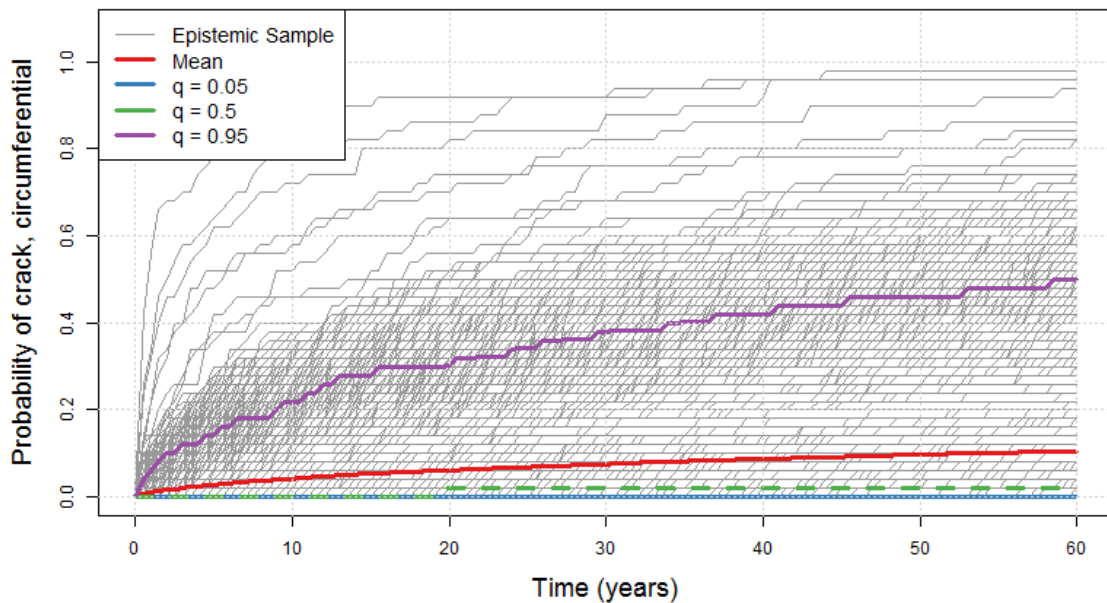


Figure 141: Probability of occurrence of circumferential crack for each epistemic realization (grey), the mean (red), and the 5th (blue), 50th (green), and 95th (purple) percentiles for Scenario 7, Run 4.

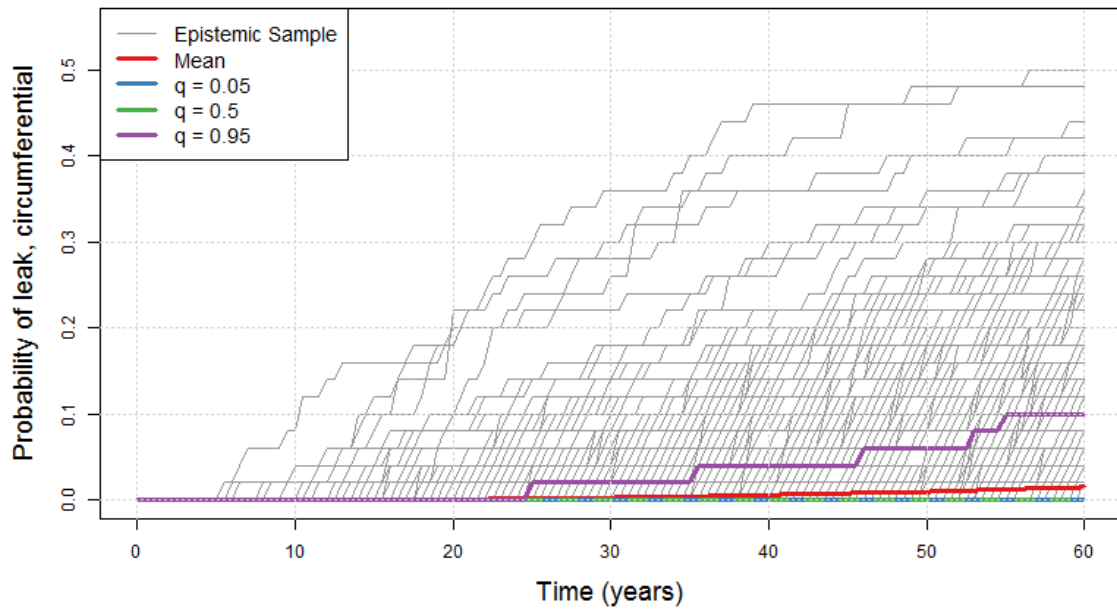


Figure 142: Probability of occurrence of circumferential leak for each epistemic realization (grey), the mean (red), and the 5th (blue), 50th (green), and 95th (purple) percentiles for Scenario 7, Run 4.

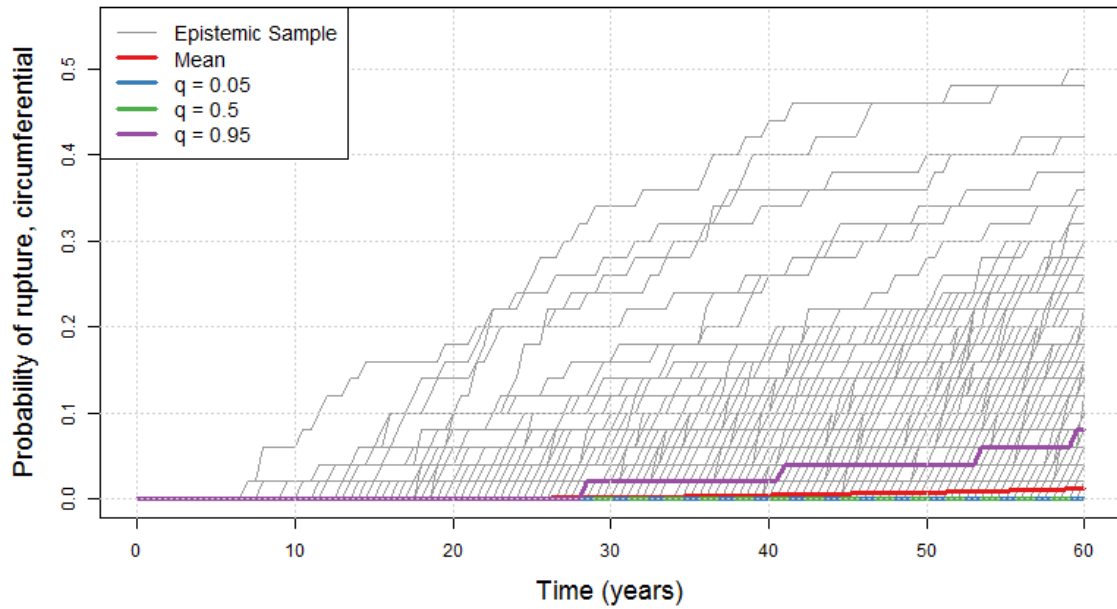


Figure 143: Probability of occurrence of circumferential rupture for each epistemic realization (grey), the mean (red), and the 5th (blue), 50th (green), and 95th (purple) percentiles for Scenario 7, Run 4.

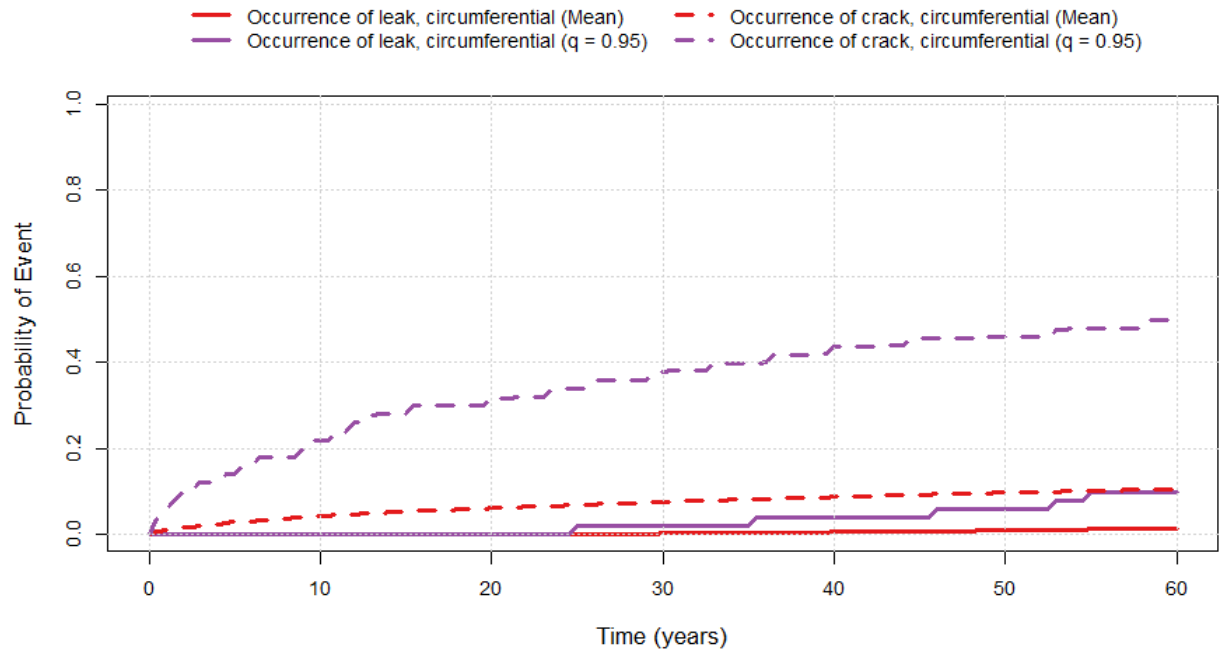


Figure 144: Mean (red) and 95th percentile (purple) for occurrence of circumferential leak (solid) and crack (dashed) for Scenario 7, Run 4.

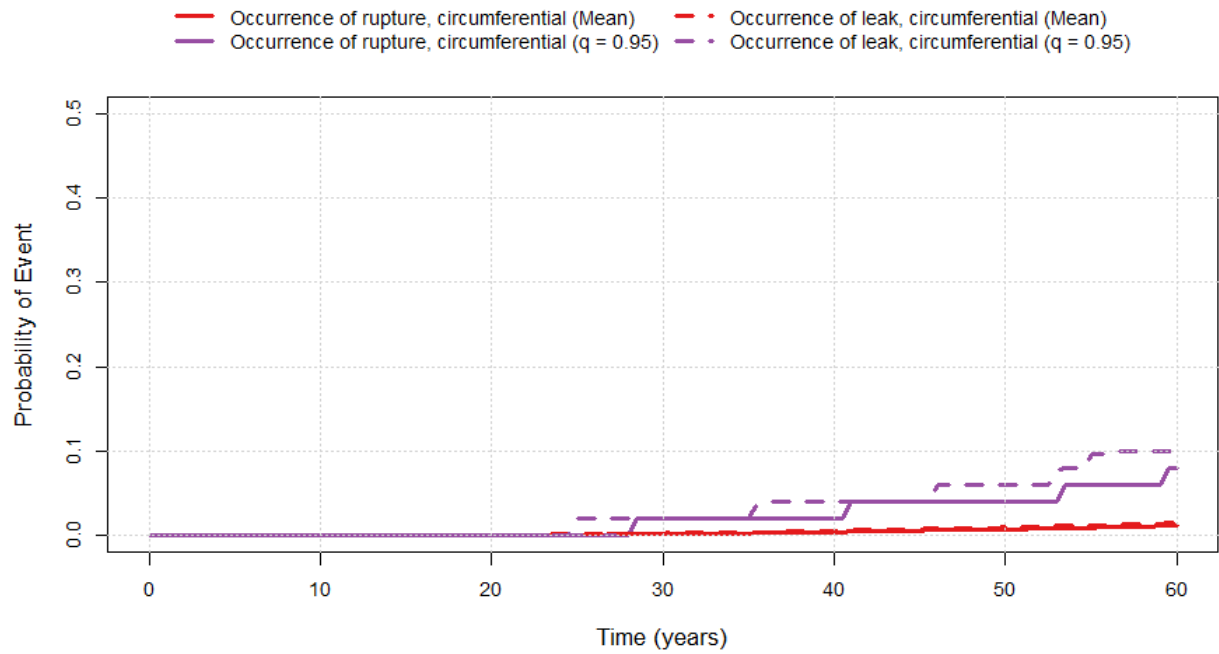


Figure 145: Mean (red) and 95th percentile (purple) for occurrence of circumferential rupture (solid) and leak (dashed) for Scenario 7, Run 4.

3.7.5 Convergence Analysis

In order to gain better insight into the convergence of the sampling options for Run 4(1000 epistemic samples with LHS and IS), this sampling scheme was re-run five times with different random number seeds for both the epistemic and aleatory loops. Each replicate was used to give an estimate of the mean probability of occurrence of circumferential ruptures. Figure 146 shows a comparison of these estimates. The replicates produce similar estimates with a range of about 0.007 at 60 years. There is no evidence of any systematic difference between the replicates for the occurrence of circumferential ruptures.

The five replicates were also used to estimate an overall mean value and a 95% prediction interval was created around this mean with a width of approximately slightly less than 0.01 at 60 years, as seen in Figure 147. This is comparable to the results from Scenario 3. This suggests that using 1000 epistemic samples with LHS and IS on p2543 produces stable estimates of the mean, regardless of the random seed used in the simulation.

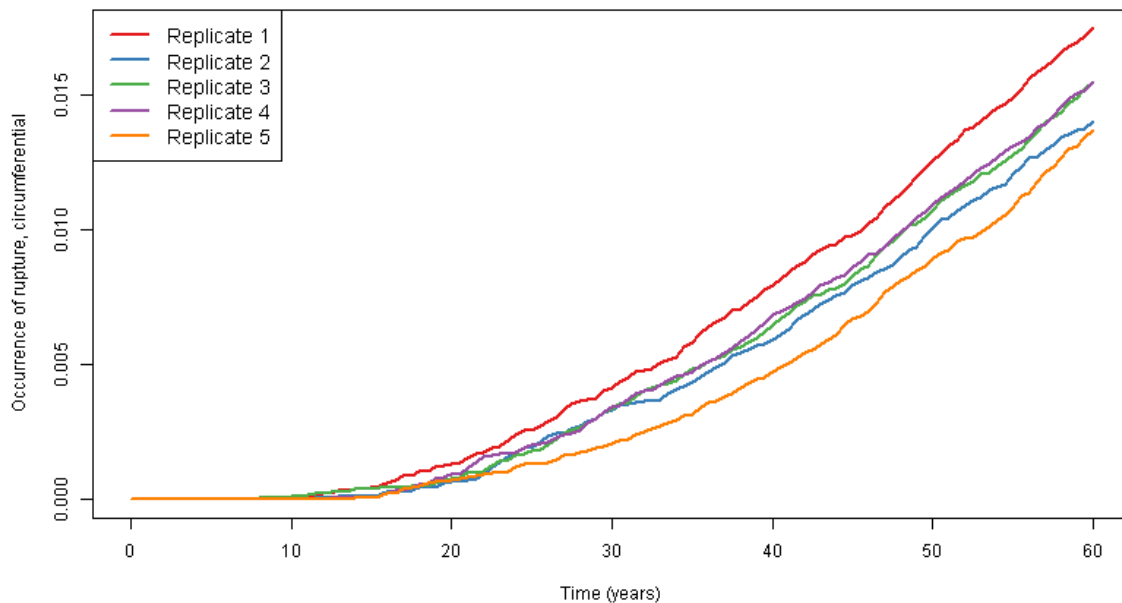


Figure 146: Comparison of estimated mean probability of occurrence of circumferential ruptures for five convergence replicates for Scenario 7, Run 4.

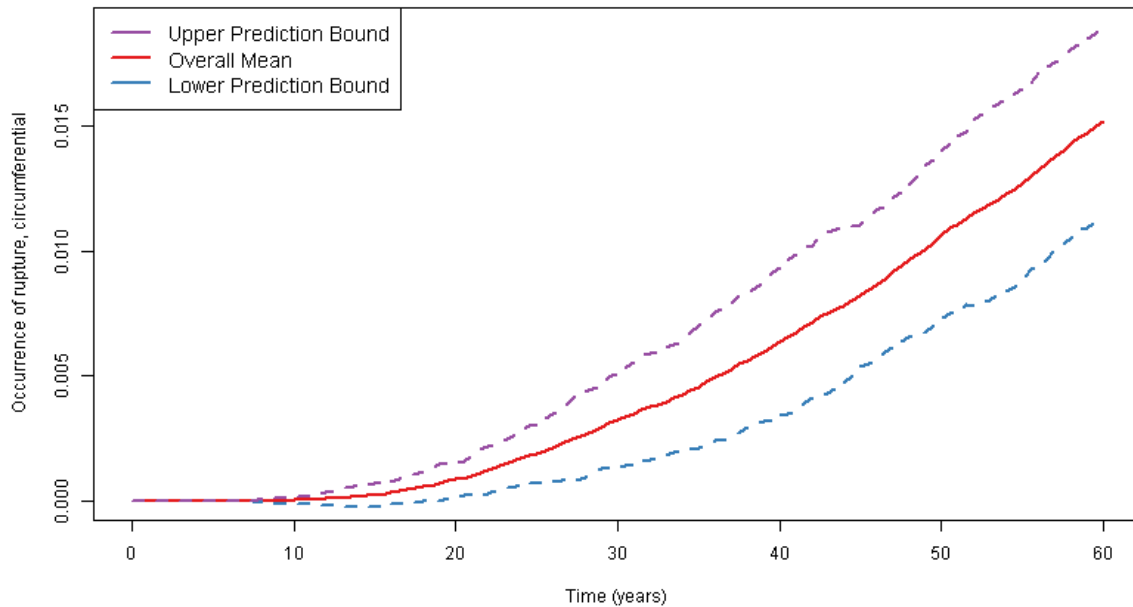


Figure 147: 95% prediction interval over mean probability of circumferential ruptures for five replicates for Scenario 7, Run 4.

3.7.6 Comparison to Scenario 3 Results

The Scenario 7 results were compared to the Scenario 3 results to examine the effect of H₂ mitigation at 20 years. The runs compared (Scenario 3 Run 9 and Scenario 7 Run 4) have the same sampling options (LHS IS 1000/50), the same epistemic and aleatory random seeds, and identical inputs except for the inclusion of H₂ mitigation at 20 years in Scenario 7. The purpose of H₂ mitigation is to reduce the probability of occurrence of leaks. This effect is illustrated in Figure 148, which compares the mean occurrences of axial leaks (dashed lines) and circumferential leaks (solid lines) between the two scenarios. The effect of the H₂ mitigation becomes apparent at 20 years and the corresponding curves continue to deviate from one another for the remainder of the simulation time. This effect is similarly demonstrated in the decreased probability of rupture after 20 years in Figure 149. The decrease in the probability of leaks for these simulations by year 60 is about 0.0075 for circumferential cracks and 0.0175 for axial cracks, which may be significant in practical applications. The reduction in leaks leads to an expected reduction in ruptures (about 0.006) by year 60, as shown in Figure 149.

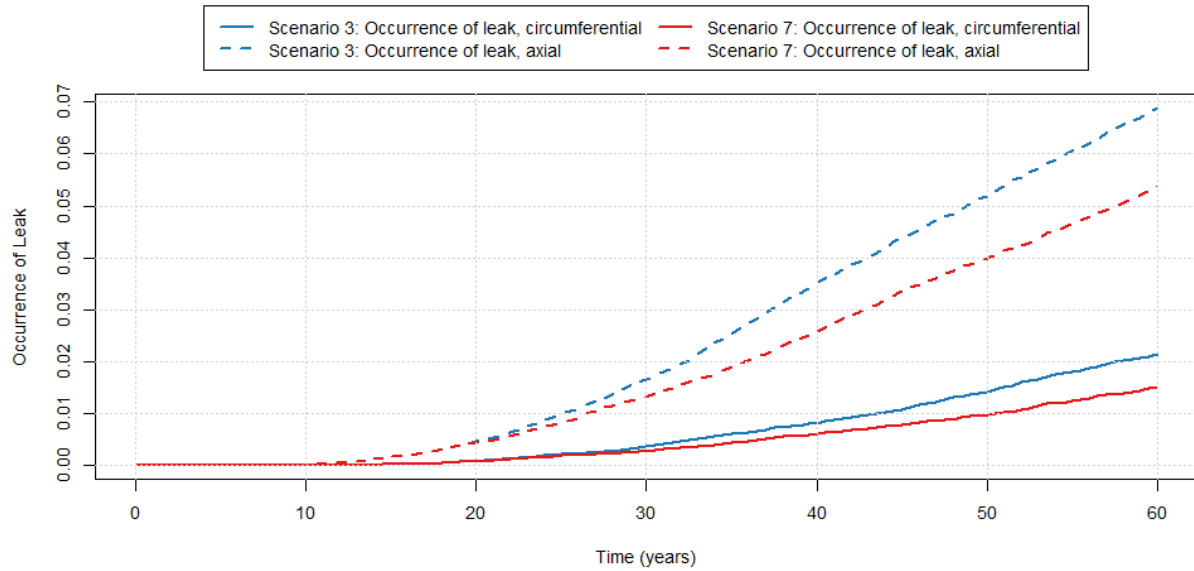


Figure 148: Mean probability of axial (dashed lines) and circumferential (solid lines) leaks occurrence for Scenario 3 (blue lines) and Scenario 7 (red lines).

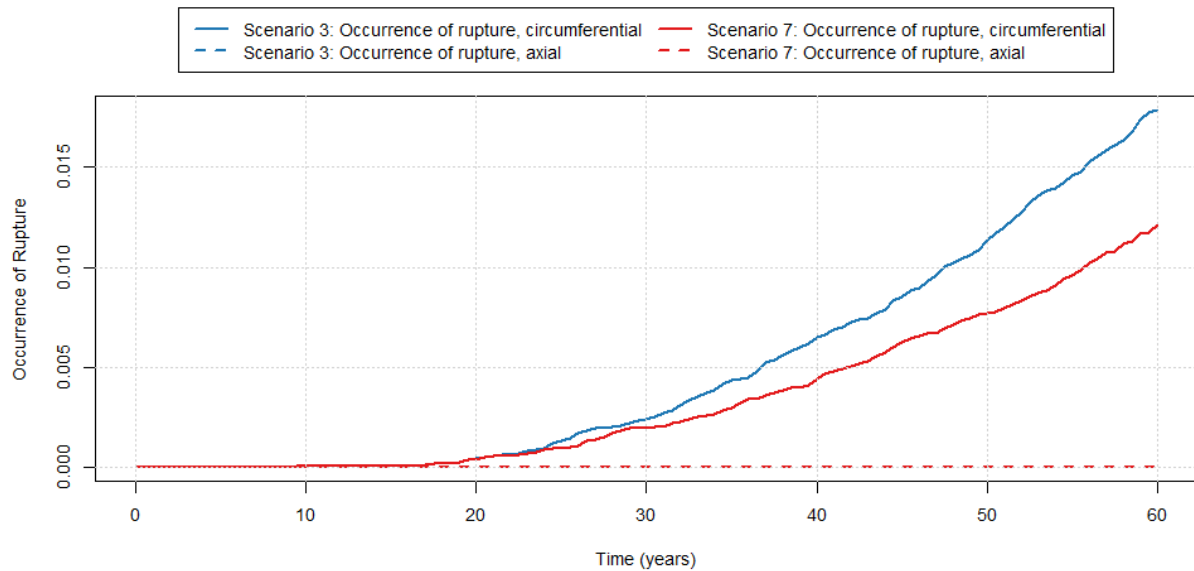


Figure 149: Mean probability of axial (dashed lines) and circumferential (solid lines) rupture occurrence for Scenario 3 (blue lines) and Scenario 7 (red lines).

The compared results for occurrence of crack between Scenario 7 and Scenario 3, shown below in Figure 150, contrast sharply to those made between Scenario 6 and Scenario 3, shown in Figure 130. As expected, the H_2 mitigation studied in Scenario 7 does not reduce the occurrence of cracks but only impacts their progression to leaks and ruptures, as described above. This contrasts to the behavior of the Zn mitigation studied in Scenario 6 which decreases the occurrence of cracks. The results for crack dimensions and leak rate similarly show a small to moderate reduction due to H_2

mitigation. The combination of these two types of chemical mitigation options was analyzed in detail for Scenario 8 with results presented in Section 3.8.

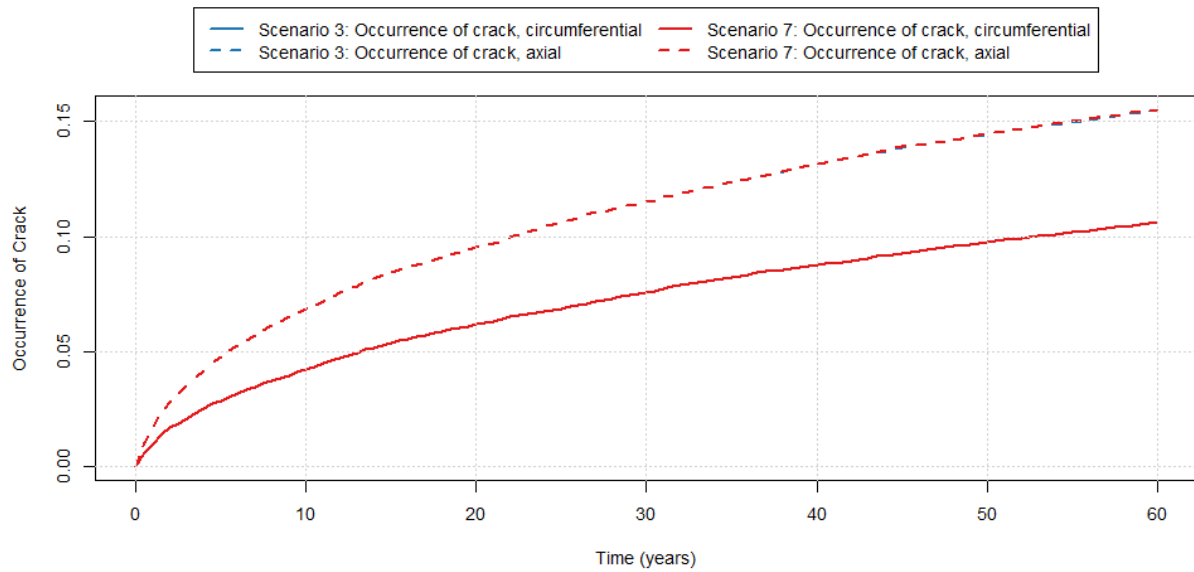


Figure 150: Mean probability of axial (dashed lines) and circumferential (solid lines) crack occurrence for Scenario 3 (blue lines) and Scenario 7 (red lines).

3.8 Scenario 8

3.8.1 Scenario Summary

The results and conclusions of a detailed analysis of sampling schemes and their impact on an understanding of the uncertainty found in the simulation results of Scenario 8 are given in this section.

Scenario 8 uses the following defining options in the xLPR code:

- PWSCC flaw initiation
- Circumferential and axial flaw orientation
- PWSCC flaw growth
- Chemical mitigation with dissolved hydrogen (H₂) and Zinc (Zn) at 20 years

This scenario is based on Scenario 3, but includes the Zn and H₂ chemical mitigations from both Scenarios 6 and 7, respectively.

The outputs of interest defined for this scenario for both circumferential and axial cracks are:

- Probability of occurrence of crack
- Probability of occurrence of leak
- Probability of occurrence of rupture
- Total Leak Rate
- Maximum crack depth
- Maximum crack inner half-length
- Maximum crack outer half-length

The comprehensive list of simulations that were used in the analysis of this scenario is shown below. User interest typically lies in understanding the uncertainty in the mean results of each output above. As a result, a constant aleatory sample size of 50 was used for each of the sampling schemes considered.

1. 100/50 (epistemic/aleatory sample size) with simple random sampling (SRS)
2. 500/50 with simple random sampling (SRS)
3. 1000/50 with Latin hypercube sampling (LHS)
4. 1000/50 with LHS and IS on p2543 with target quantile of 0.95
5. 1000/50 with SRS and IS on p2543 with target quantile of 0.95

Schemes were chosen based on insight gained from the results of the Scenario 3 analysis; those schemes that were most informative and exercised the epistemic sampling options in the code were chosen for Scenario 8.

The results of the Scenario 8 analysis are consistent with the analysis of Scenario 3. The only difference between the scenarios is that Scenario 8 includes H₂ and Zn mitigation at 20 years; in every other respect the scenario is identical to Scenario 3. The results of regression analyses indicate that the multiplier on Direct Model 1 (DM1) proportionality constant A for the PWSCC crack initiation module, hereafter referred to as p2543, was the most important input for explaining

variability in all of the outputs for Scenario 8. This is consistent with the results from Scenarios 6 and 7, which respectively include Zn and H₂ mitigation separately. When used, importance sampling was applied to p2543 to improve the precision in the output of interest.

As in the Scenario 3 analysis, overall, results confirm that increasing the epistemic sample size decreases the uncertainty in the mean of the output (e.g., probability of occurrence of leak from axial cracks) as expected, especially when moving from 100 to 500 epistemic samples. The decrease in uncertainty in the mean of the output is also evident when transitioning from 500 to 1000 epistemic realizations. An epistemic sample size of 1000 is sufficient for this scenario, though sample size ultimately depends upon user needs.

3.8.2 Analysis Progression

The five sampling schemes, informed by the Scenario 3 analysis in Section 3.3, were chosen to demonstrate stability and the impact of sampling options on the stability of results. These schemes progress in sample size and sampling options.

Runs 1 and 2 are identical in all sampling options except the epistemic sample size. The epistemic sample size for Run 1 was 100 and this was increased to 500 for Run 2. The mean estimate of the probability of circumferential rupture from Runs 1 and 2 is plotted in Figure 151 with 95% confidence intervals calculated using the bootstrap method. Analysis of this particular output is illustrative of the effects of both Zn and H₂ mitigation. Zn mitigation prevents or postpones the initiation of cracks, leading cracks to form later in the simulation so they do not grow as large within the simulation time, leading to a lower leak rate and, ultimately, fewer ruptures. H₂ mitigation should slow the growth of cracks, leading to an overall decrease in crack size and ruptures. The mean estimates for Run 1 and Run 2 are close initially, but deviate more starting at about 20 years. This may reflect the effect of mitigation, with the deviation suggesting that larger sample sizes are needed to better approximate the reduction in crack initiation and growth, and hence in the probability of rupture, when mitigation is used.

With regard to stability, the width of the confidence intervals around the estimated mean decreases as expected when the epistemic sample size is increased to 500. This is shown in Figure 152. Whether this decrease is significant is application specific; with 100 epistemic samples, the estimated probability of circumferential rupture at year 60 covers a range with width about 0.0155. Increasing the sample size to 500 reduces the width of this interval to about 0.0058.

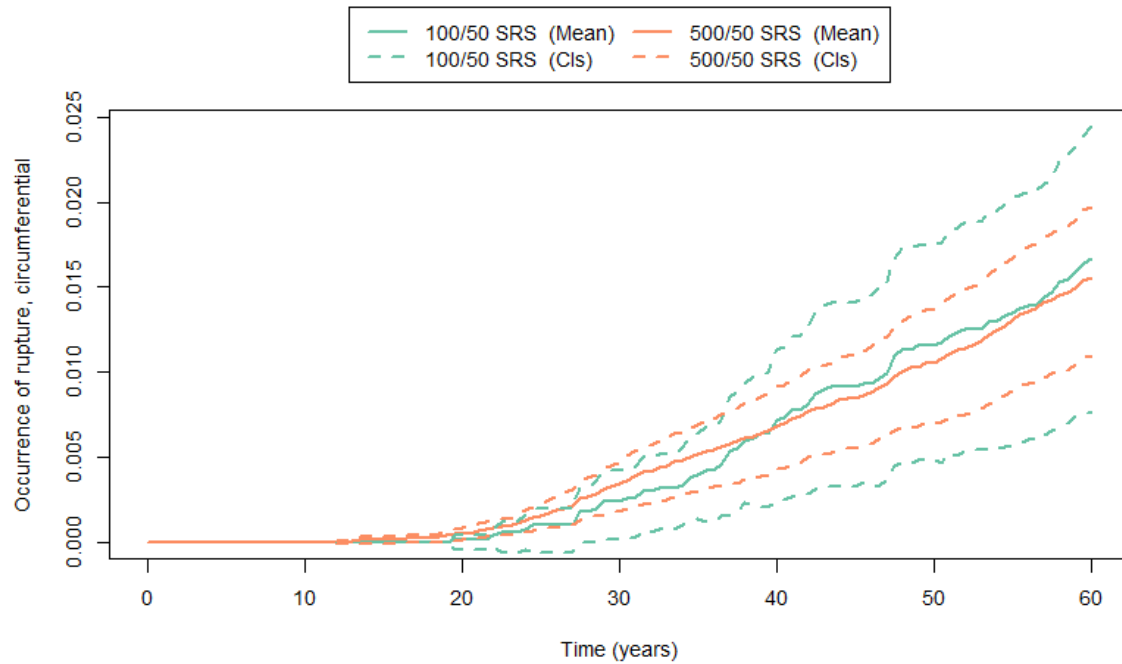


Figure 151: Mean estimate of the probability of circumferential ruptures (solid lines) and 95% confidence intervals (dashed lines) for Scenario 8, Run 1 (green) and Run 2 (orange).

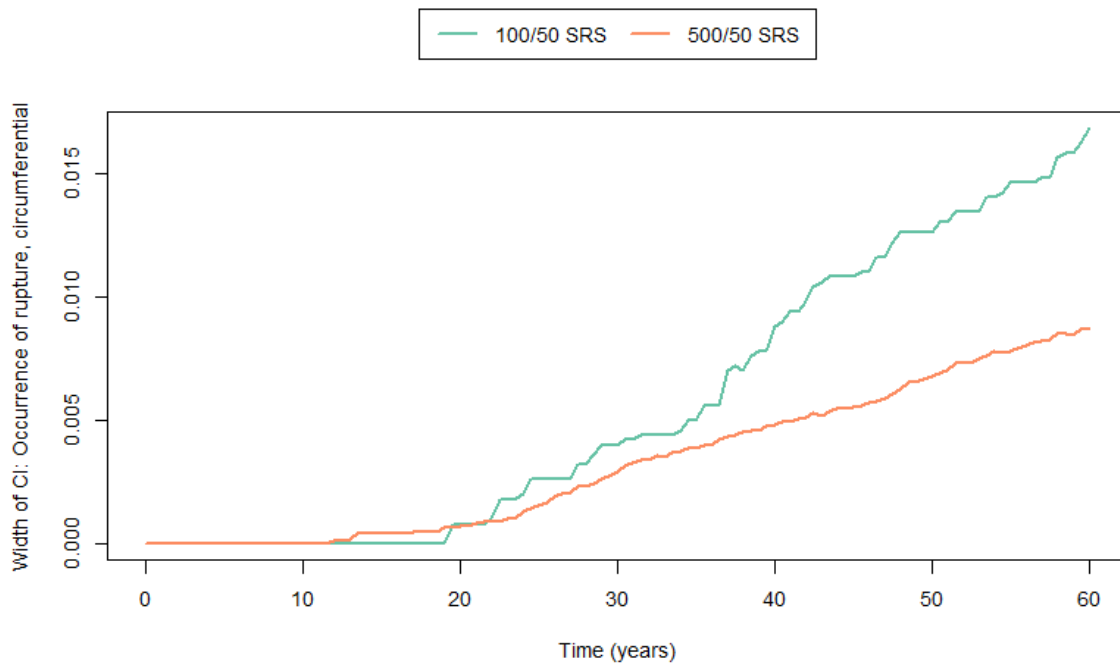


Figure 152: Width of 95% confidence interval for the estimate of mean probability of circumferential ruptures for Scenario 8, Run 1 (green) and 2 (orange).

The relatively large decrease in confidence interval width between Runs 1 and 2 compared to the magnitude of the estimate suggest that an increase in samples, and thus simulation time, may still be worth the benefit in reducing uncertainty in the final mean estimate. The Scenario 3 analysis showed that LHS sampling could improve stability and there is no computational reason to prefer SRS to LHS. Thus, to improve upon the results in Run 2, the epistemic sample size was increased to 1000 and LHS was used instead of SRS to sample epistemically uncertain parameters for Scenario 8.

The increase in sample size and change in sampling strategy from Run 2 to Run 3 decreased the width of the confidence interval, shown in Figure 153 and Figure 154, by about one-half the decrease from Run 1 to Run 2. However, the mean estimate of the probability of circumferential rupture is substantially lower when LHS is employed compared to SRS; the mean estimate from Run 2 is not contained in the confidence intervals around the mean estimate from Run 3 starting at about 25 years. This motivates the inclusion of IS to examine stability.

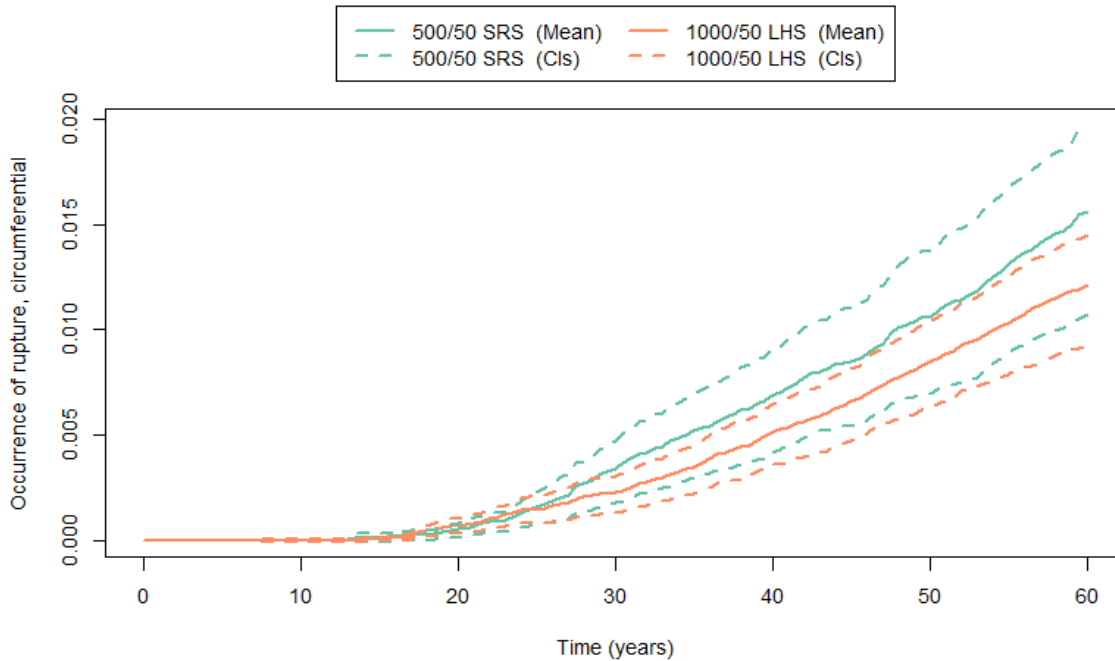


Figure 153: Mean estimate of the probability of circumferential ruptures (solid lines) and 95% confidence intervals (dashed lines) for Scenario 8, Run 2 (green) and Run 3 (orange).

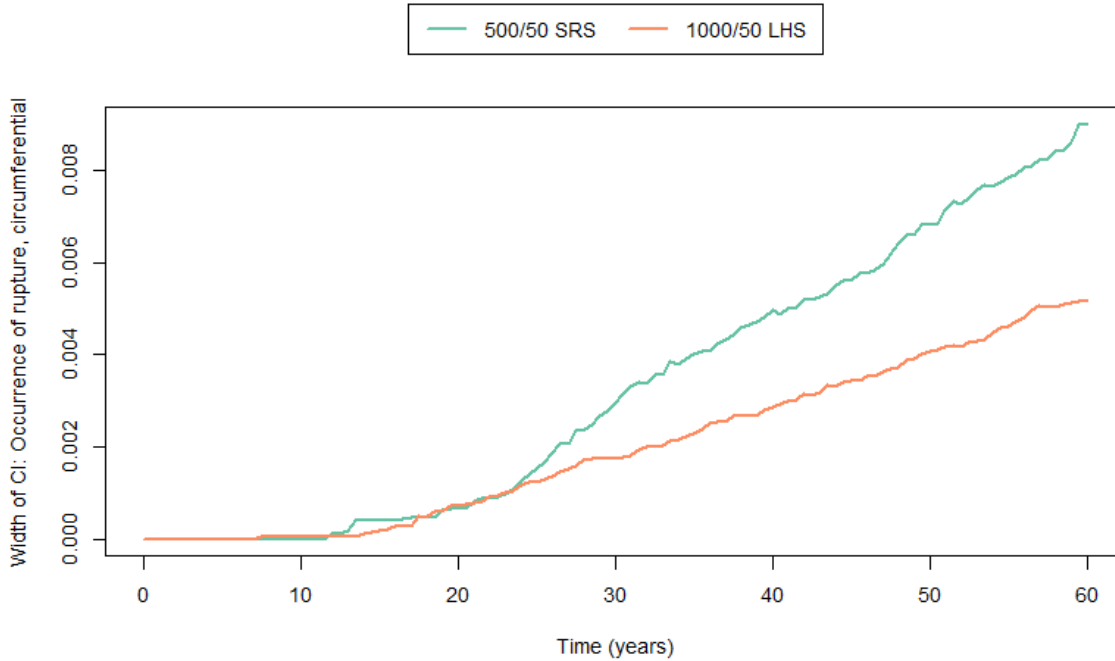


Figure 154: Width of the 95% confidence interval for the mean probability of circumferential ruptures for Scenario 8, Run 2 (green) and Run 3 (orange).

Stepwise rank regression results from Run 3 on the total leak rate for Scenario 8 are presented in Table 9. These results are consistent with regression results from Runs 1 and 2. The parameter identified in Runs 1, 2, and 3 of Scenario 8 as the most important variable with respect to variability in the occurrence of circumferential ruptures is p2543, the multiplier on DM1 proportionality constant A. The influence of this variable can be seen both by the high SRRC value (0.521) and the scatterplots of p2543 with the probability of occurrence of circumferential ruptures for the first three runs in Figure 155. This is consistent with the conclusion for all of the other outputs of interest.

Table 9: Summary of the stepwise rank regression analysis results for all outputs at year 60 for Scenario 8, Run 3.

Variable Identifier	Variable Name	SRRC	Variable Identifier	Variable Name	SRRC
Probability of Occurrence of Axial Cracks – $R^2 = 0.862$			Maximum Axial Crack Depth – $R^2 = 0.800$		
p2543	Multiplier proport. Const. A (DM1)	0.909	p2543	Multiplier proport. Const. A (DM1)	0.850
p4350	Hoop WRS Pre-mitigation	-0.187	p4350	Hoop WRS Pre-mitigation	-0.231
p5106	b (circ)	0.064	p2592	Comp-to-Comp Variab Factor, fcomp	0.083
p2531	Zn Factor of Improvement - 1, FOIZn-1	-0.056	p1102	Pipe Wall Thickness	-0.082
p1102	Pipe Wall Thickness	-0.049	p2595	Charact Width of Peak vs ECP, c	0.081
Probability of Occurrence of Circumferential Cracks – $R^2 = 0.809$			Maximum Circumferential Crack Depth – $R^2 = 0.780$		
p2543	Multiplier proport. Const. A (DM1)	0.891	p2543	Multiplier proport. Const. A (DM1)	0.859
p4352	Axial WRS Pre-mitigation	-0.116	p4352	Axial WRS Pre-mitigation	-0.153
p2531	Zn Factor of Improvement - 1, FOIZn-1	-0.071	p2592	Comp-to-Comp Variab Factor, fcomp	0.068
p1102	Pipe Wall Thickness	-0.022	p2594	Peak-to-Valley ECP Ratio - 1, P-1	-0.065
p5108	b (axial)	0.021	p1102	Pipe Wall Thickness	-0.059
Probability of Axial Leak – $R^2 = 0.606$			Maximum Axial Inner Half-Length – $R^2 = 0.851$		
p2543	Multiplier proport. Const. A (DM1)	0.698	p2543	Multiplier proport. Const. A (DM1)	0.897
p4350	Hoop WRS Pre-mitigation	-0.201	p4350	Hoop WRS Pre-mitigation	-0.218
p2592	Comp-to-Comp Variab Factor, fcomp	0.194	p5106	b (circ)	0.067
p2595	Charact Width of Peak vs ECP, c	0.178	p1102	Pipe Wall Thickness	-0.056
p2594	Peak-to-Valley ECP Ratio - 1, P-1	-0.144	p2531	Zn Factor of Improvement - 1, FOIZn-1	-0.055
Probability of Circumferential Leak – $R^2 = 0.437$			Maximum Circumferential Inner Half-Length – $R^2 = 0.764$		
p2543	Multiplier proport. Const. A (DM1)	0.563	p2543	Multiplier proport. Const. A (DM1)	0.845
p2592	Comp-to-Comp Variab Factor, fcomp	0.182	p4352	Axial WRS Pre-mitigation	-0.157
p4352	Axial WRS Pre-mitigation	-0.178	p2592	Comp-to-Comp Variab Factor, fcomp	0.091
p1102	Pipe Wall Thickness	-0.153	p2594	Peak-to-Valley ECP Ratio - 1, P-1	-0.079
p2594	Peak-to-Valley ECP Ratio - 1, P-1	-0.144	p2595	Charact Width of Peak vs ECP, c	0.075
Probability of Circumferential Rupture – $R^2 = 0.411$			Maximum Axial Outer Half-Length – $R^2 = 0.605$		
p2543	Multiplier proport. Const. A (DM1)	0.521	p2543	Multiplier proport. Const. A (DM1)	0.697
p2592	Comp-to-Comp Variab Factor, fcomp	0.208	p4350	Hoop WRS Pre-mitigation	-0.201
p1102	Pipe Wall Thickness	-0.181	p2592	Comp-to-Comp Variab Factor, fcomp	0.195
p2594	Peak-to-Valley ECP Ratio - 1, P-1	-0.167	p2595	Charact Width of Peak vs ECP, c	0.177
p4352	Axial WRS Pre-mitigation	-0.163	p2594	Peak-to-Valley ECP Ratio - 1, P-1	-0.144
Total Leak Rate – $R^2 = 0.597$			Maximum Circumferential Outer Half-Length – $R^2 = 0.436$		
p2543	Multiplier proport. Const. A (DM1)	0.696	p2543	Multiplier proport. Const. A (DM1)	0.561
p2592	Comp-to-Comp Variab Factor, fcomp	0.201	p2592	Comp-to-Comp Variab Factor, fcomp	0.187
p2595	Charact Width of Peak vs ECP, c	0.176	p4352	Axial WRS Pre-mitigation	-0.178
p2594	Peak-to-Valley ECP Ratio - 1, P-1	-0.156	p2594	Peak-to-Valley ECP Ratio - 1, P-1	-0.150
p1102	Pipe Wall Thickness	-0.119	p1102	Pipe Wall Thickness	-0.149

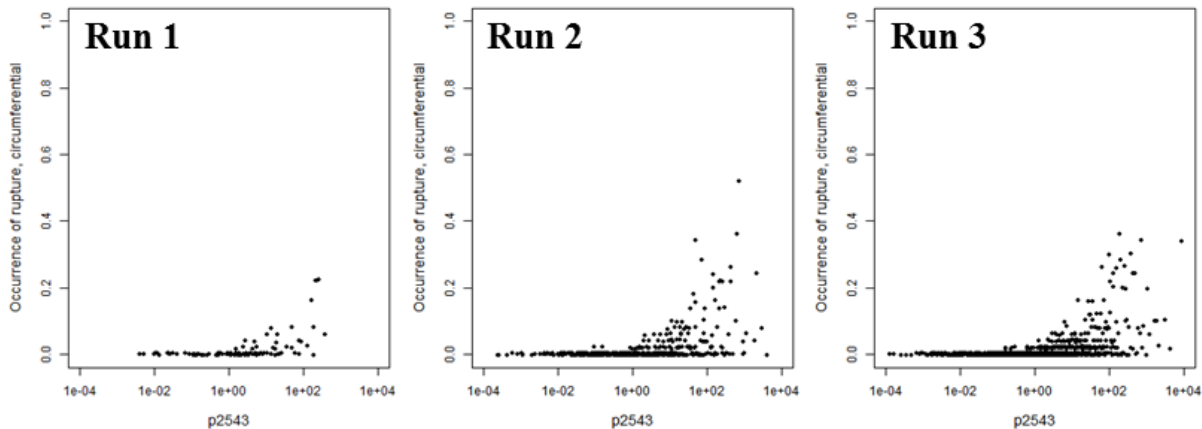


Figure 155: Scatter plots for variable p2543 (multiplier on DM1 proportionality constant A) and the occurrence of rupture for Scenario 8, Run 1 (left), Run 2 (center), and Run 3 (right).

Run 4 expands on Run 3 by incorporating importance sampling, while maintaining the same sample size (1000) and sampling strategy (LHS). Importance sampling was applied on p2543 at the 95th quantile to oversample the upper tail of the p2543 distribution to increase the probability of circumferential rupture (in particular by increasing the occurrence of cracks). The comparison between mean estimates of the probability of circumferential rupture for Runs 3 and 4 is plotted in Figure 156, showing a potentially significant effect from including IS. This suggests that, while this level of convergence might be acceptable for some applications, other applications may require larger simulations.

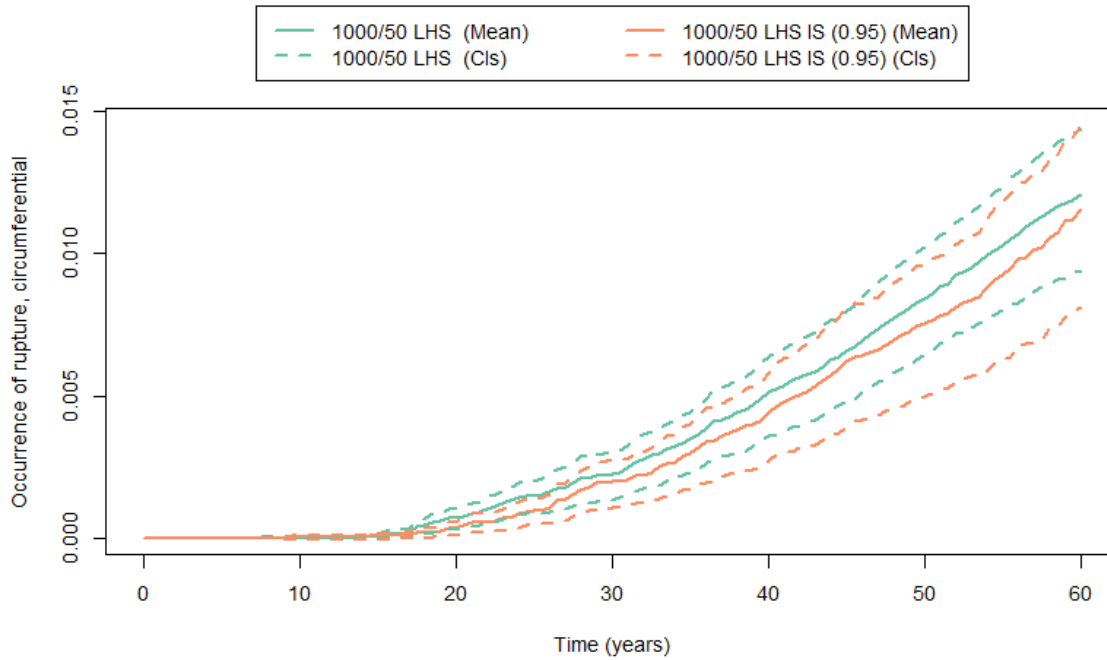


Figure 156: Mean probability of occurrence of circumferential ruptures (solid line) and 95% confidence intervals (dashed lines) for Scenario 8, Run 3 (green) and Run 4 (orange).

Stability is also questionable based on the results from Run 5 which maintains the sample size and importance sampling on p2543 from Run 4, but employs SRS rather than LHS. The results, in Figure 157, appear to be significantly different, with each mean contained only in its own confidence interval in the latter half of the simulation. This level of instability does not occur with all of the outputs. Notice the plots with mean estimate of probability of occurrence of circumferential crack and of circumferential leak with confidence intervals for Run 4 and Run 5 in Figure 158 and Figure 159 respectively. Figure 158 supports convergence for the estimate of mean probability of occurrence of circumferential cracks, even though the estimate of probability of circumferential ruptures may not have converged. Figure 159, however, shows a similar lack of convergence for the probability of occurrence of circumferential leak as for circumferential rupture. Note, however, that the probabilities of occurrence of circumferential leak and rupture are estimated to be roughly an order of magnitude lower than the probability of crack. This supports the effectiveness of mitigation but also makes it more difficult to estimate the probabilities for leak and rupture.

Together, these plots suggest that with SRS and LHS, the low probability of occurrence of rupture when both chemical mitigation types are included may be why the probability of occurrence of circumferential rupture converges slowly; with fewer leaks and younger cracks, there are fewer rupture events to use in estimating the probability of a rupture. Importance sampling is intended to solve precisely this problem. However, the effectiveness of importance sampling is dependent upon the quantile used. If the level of uncertainty around the estimated probability of circumferential rupture is too high within context, this is an indication that a different importance sampling quantile might be necessary.

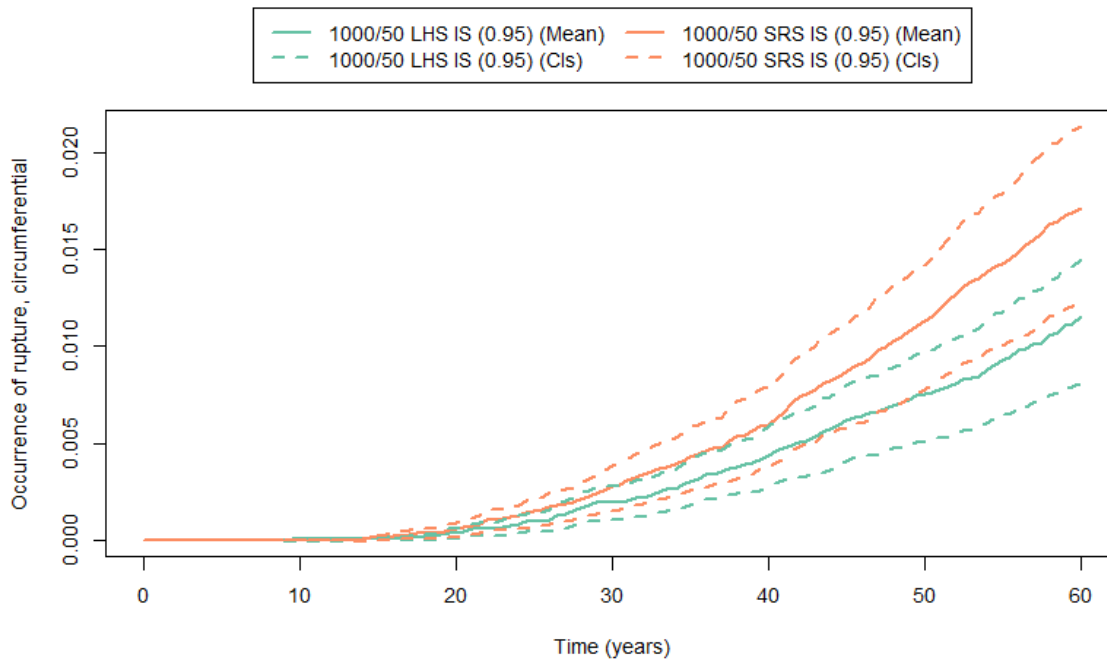


Figure 157: Mean probability of occurrence of circumferential ruptures (solid line) and 95% confidence intervals (dashed lines) for Scenario 8, Run 4 (green) and Run 5 (orange).

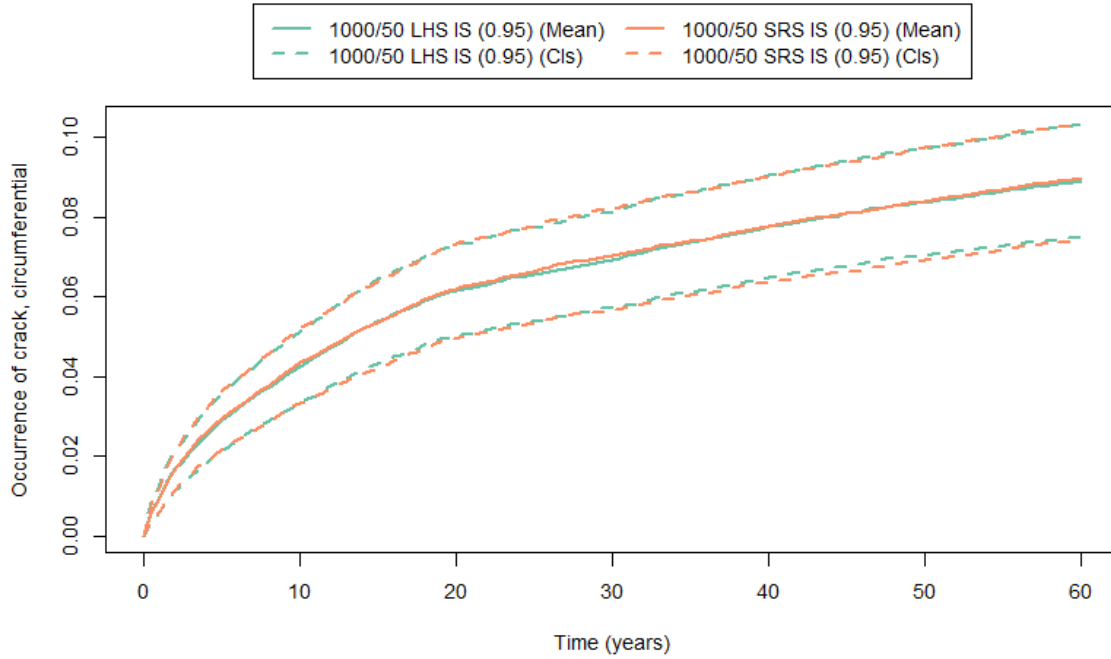


Figure 158: Mean probability of occurrence of circumferential cracks (solid line) and 95% confidence intervals (dashed lines) for Scenario 8, Run 4 (green) and Run 5 (orange).

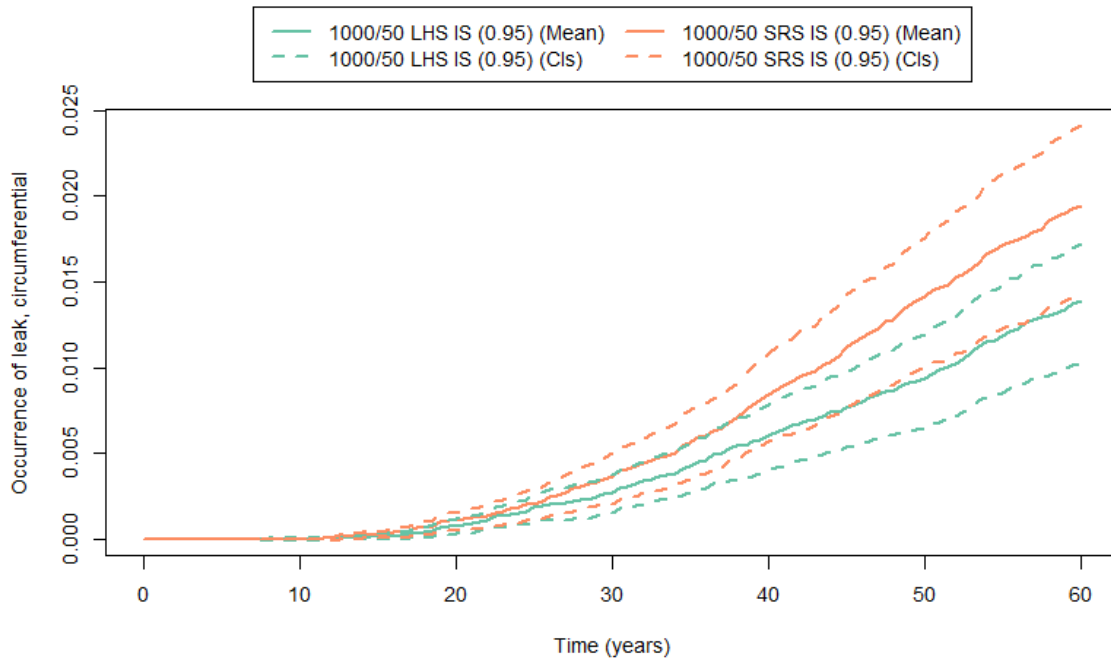


Figure 159: Mean probability of occurrence of circumferential leaks (solid line) and 95% confidence intervals (dashed lines) for Scenario 8, Run 4 (green) and Run 5 (orange).

3.8.3 Overall Results

The iterative process and run-to-run comparisons described in the previous section resulted in five runs that characterize the stability of simulation results with respect to sampling options. A comparison across this collection of runs can be used to provide guidance on choosing a sampling scheme for Scenario 8.

An ideal selection of sampling options leads to modeling results that converge within an acceptable computational time. Figure 160 shows the results for the probability of circumferential rupture for Runs 1 through 5. Though the results may appear to suggest that sampling schemes using SRS with no IS behave differently than the rest of the runs, this is due to the random seed choice and does not indicate a tendency for SRS to over-estimate generally, as studied in-depth in Section 4.4. However, these results suggest that more advanced sampling methods (LHS or LHS with IS) may be particularly advantageous in this scenario.

All of the other sampling schemes lead to results that are approximately equal. These results show that, while comparing Runs 3, 4, and 5 pairwise can suggest a lack of convergence, there is also a clear trend common to these runs. If the order of the uncertainty is sufficiently small for the application, these runs can be considered sufficiently converged.

The uncertainty associated with the result of interest, shown in the form of 95% confidence intervals, can also be used to identify beneficial sampling selections for this scenario. Figure 152 and Figure 154 show how uncertainty is affected by the epistemic sample size. Choosing the final sample size will depend on both the computational time available and the level of uncertainty that the user is willing to accept. A smaller sample size may be more reasonable if an output that converges faster, such as the probability of occurrence of circumferential cracks, is the primary motivation for the analysis.

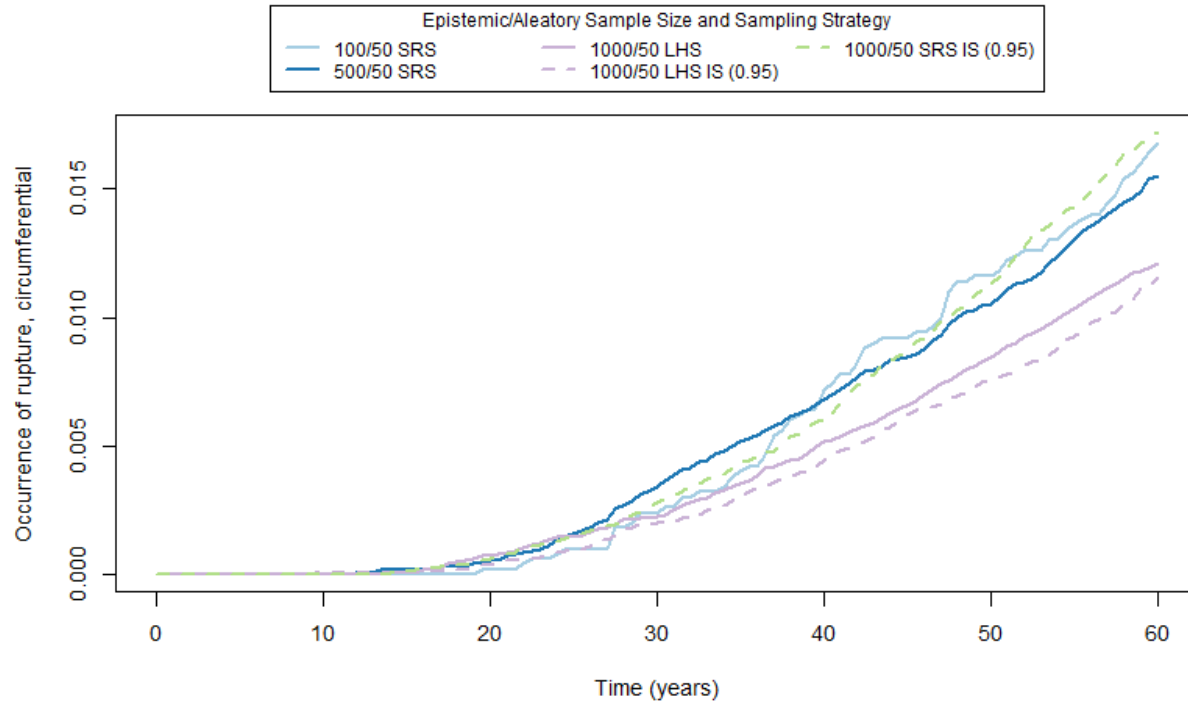


Figure 160: Mean estimate of the probability of occurrence of circumferential ruptures for Scenario 8, Runs 1 through 5.

Importance sampling had an impact on decreasing the uncertainty in the result of interest for runs using SRS, LHS, and for runs with large epistemic sample sizes. For outputs like the probability of occurrence of crack, importance has a less significant impact, particularly when a large sample size and LHS are used. For outputs that converge more slowly, like the probability of occurrence of circumferential rupture, importance sampling is likely necessary to reduce uncertainty significantly. This is a departure from the results of Scenario 3 may be because the combined effects of Zn and H₂ mitigation lead to a significant decrease in events (cracks, leaks, ruptures) that cannot be wholly overcome with importance sampling alone.

3.8.4 Epistemic Uncertainty Analysis

An uncertainty analysis was conducted for Run 4 with 1000 epistemic and 50 aleatory sample with LHS and IS for Scenario 8. Plots of the estimated probabilities of occurrence of circumferential crack, leak, and rupture are displayed in Figure 161, Figure 162, and Figure 163, respectively. These figures show the probability of occurrence for each epistemic realization (grey lines), the mean (red line), and the 5th, 50th, and 95th percentiles (blue, green, and purple lines, respectively). In contrast to the 95% confidence intervals for the mean shown in Section 3.8.2, which were calculated using the bootstrap method, the percentiles shown here offer a summary of the total range of results. They encompass 90% of all the epistemic realizations observed in this run. For example, the 50th percentile for circumferential crack is about 0.02 at year 60, meaning that half of the epistemic realizations have at least a 2% chance of a crack by year 60. For leak and rupture, the 5th and 50th percentiles are uniformly zero up to year 60, indicating that at least 50% of the realizations did not experience a circumferential leak or rupture in the first 60 years. These results are consistent with

the results from Scenario 3, showing a slightly decreased probability of circumferential cracks and moderately decreased probabilities of leak and rupture by year 60 with H₂ and Zn mitigation at 20 years.

As seen in the Scenario 3 results, the epistemic realizations for leak are also shifted to the right in Scenario 8 compared to cracks and the epistemic realizations for rupture are also shifted to the right compared to leaks. This result indicates that for this scenario, cracks are delayed but do progress to leaks and eventually ruptures; H₂ mitigation may slow the growth of cracks enough to significantly delay leaks. This relationship is further illustrated in Figure 164 and Figure 165, which display the mean and 95th percentile for crack and leak and for leak and rupture, respectively. These plots also show that the spread between the mean and the 95th percentile is larger for crack than it is for leak and larger for leak than it is for rupture— indicating rupture and leak results are relatively more certain. This is somewhat intuitive based on the large spread in the probability of occurrence of circumferential cracks in Figure 161 compared to those for occurrence of leaks in Figure 162 and those for occurrence of ruptures in Figure 163. The probabilities of leak and rupture and also smaller, due to mitigation, and thus more difficult to estimate. Extreme quantiles are more difficult to estimate than the mean, so if the user is concerned with estimating the 95th percentile, a larger sample size may be necessary.

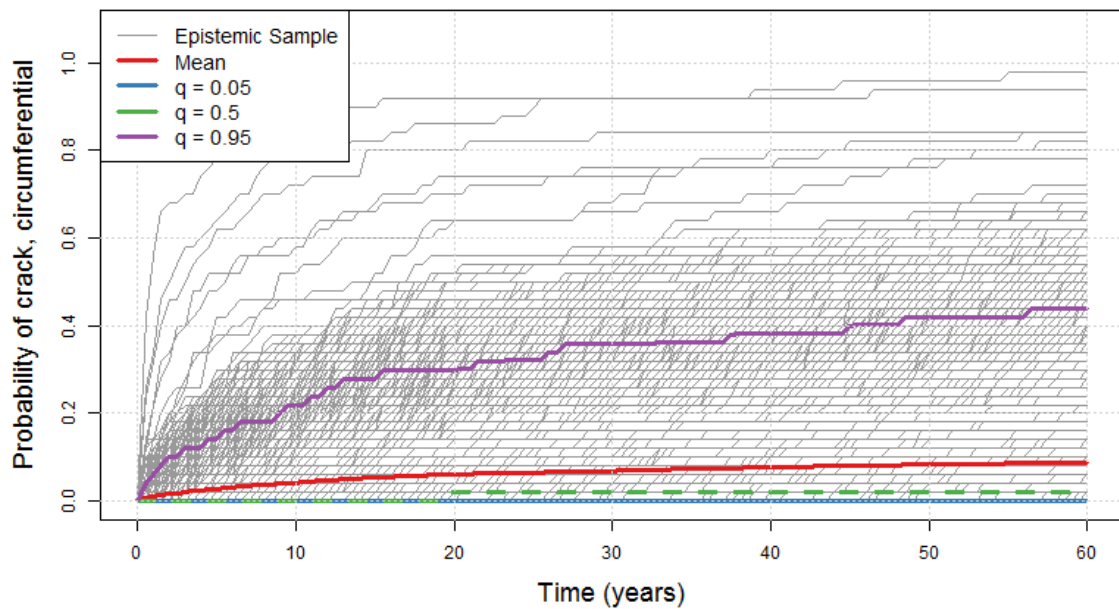


Figure 161: Probability of occurrence of circumferential crack for each epistemic realization (grey), the mean (red), and the 5th (blue), 50th (green), and 95th (purple) percentiles for Scenario 8, Run 4.

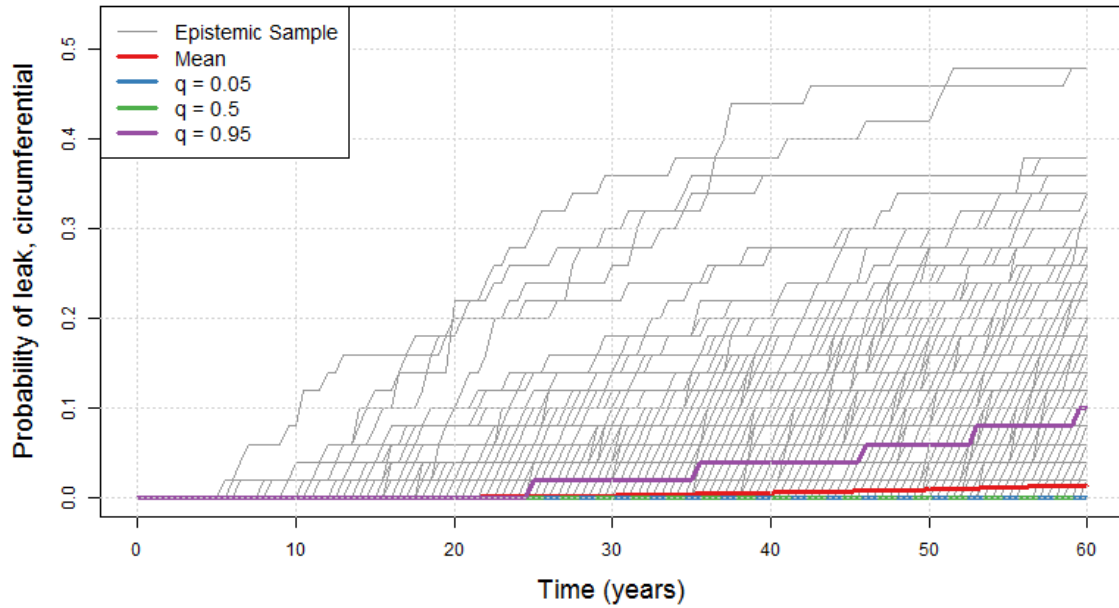


Figure 162: Probability of occurrence of circumferential leak for each epistemic realization (grey), the mean (red), and the 5th (blue), 50th (green), and 95th (purple) percentiles for Scenario 8, Run 4.

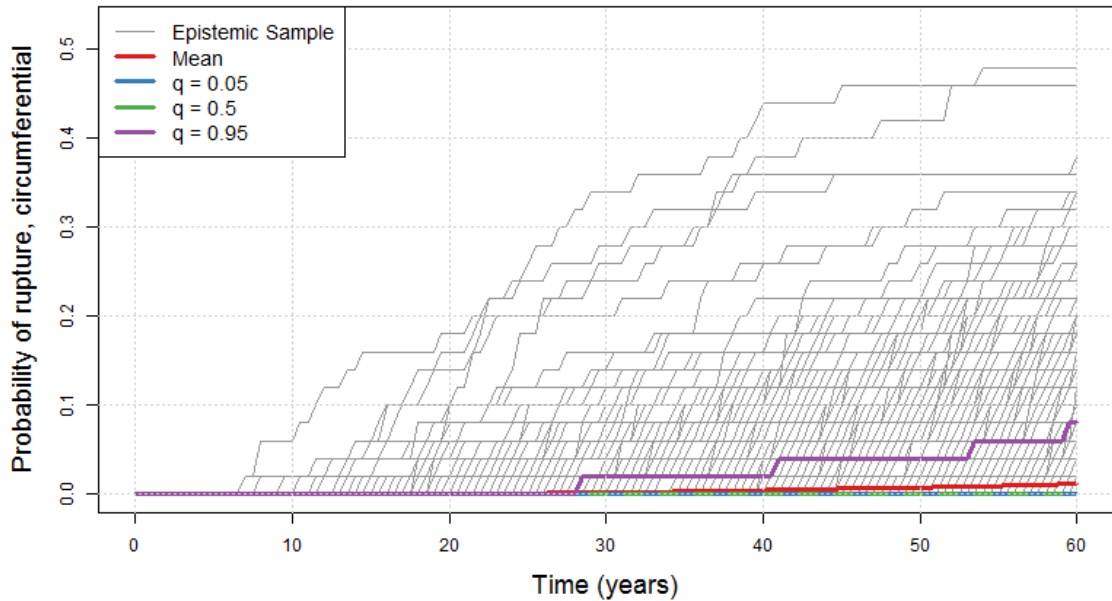


Figure 163: Probability of occurrence of circumferential rupture for each epistemic realization (grey), the mean (red), and the 5th (blue), 50th (green), and 95th (purple) percentiles for Scenario 8, Run 4.

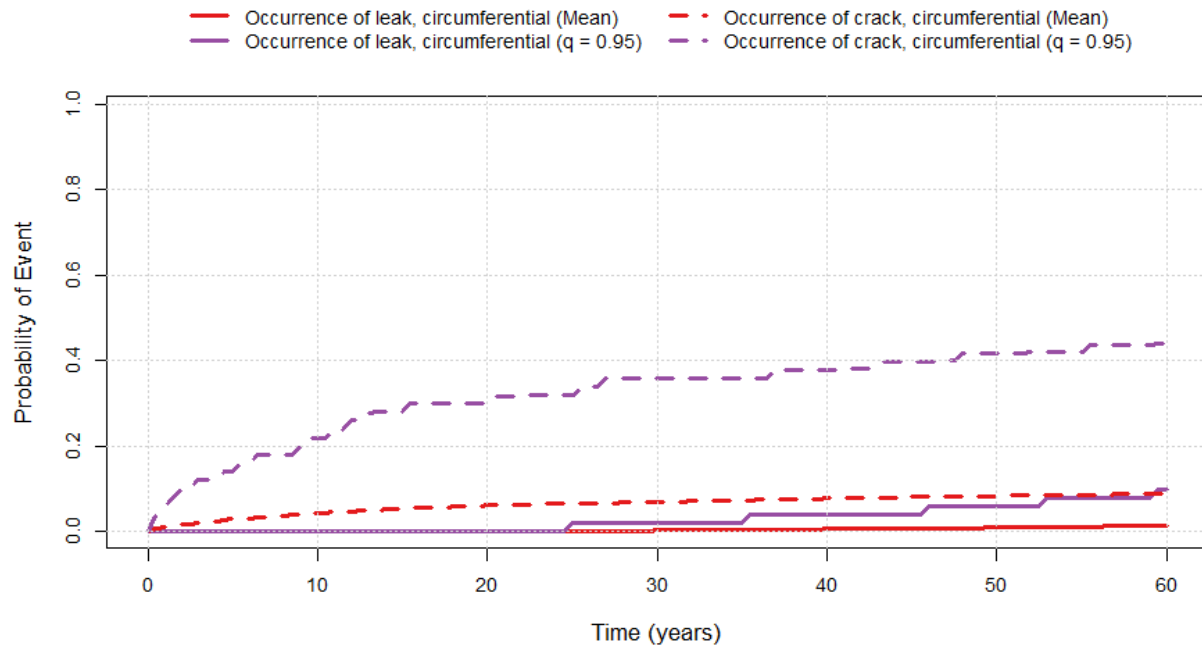


Figure 164: Mean (red) and 95th percentile (purple) for occurrence of circumferential leak (solid) and crack (dashed) for Scenario 8, Run 4.

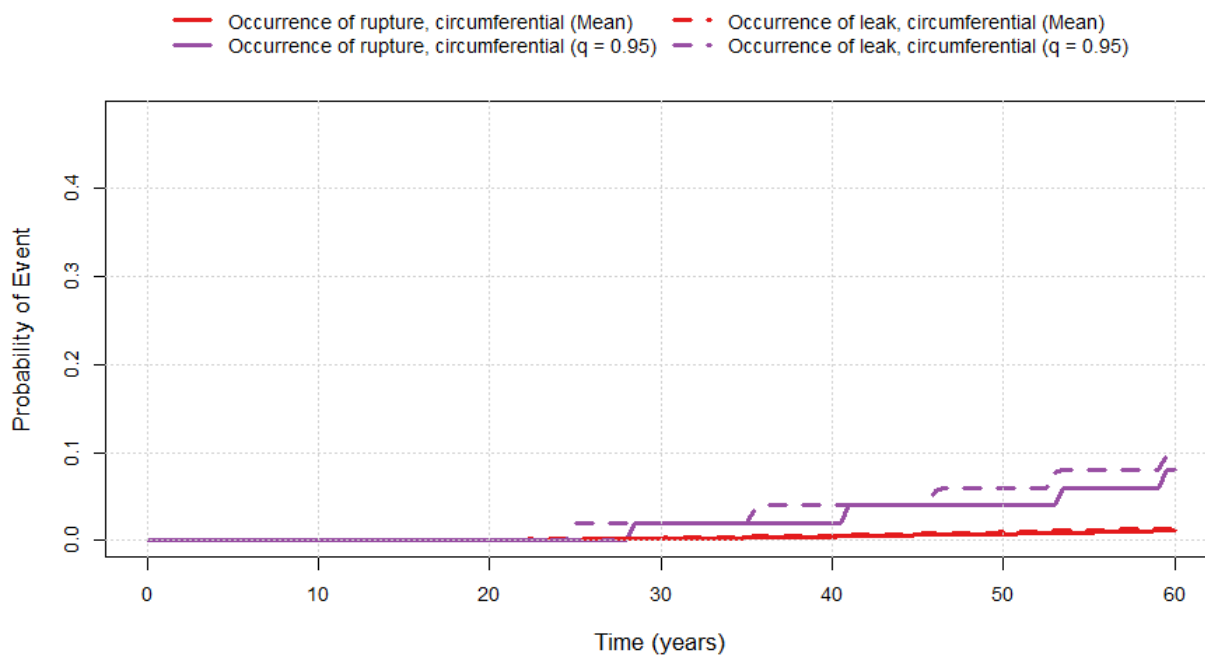


Figure 165: Mean (red) and 95th percentile (purple) for occurrence of circumferential rupture (solid) and leak (dashed) for Scenario 8, Run 4.

3.8.5 Convergence Analysis

In order to gain better insight into the convergence of the sampling options for Run 4 (1000 epistemic samples with LHS and IS), this sampling scheme was re-run five times with different random number seeds for both the epistemic and aleatory loops. Each replicate was used to give an estimate of the mean probability of occurrence of circumferential cracks and the maximum inner half-length for circumferential cracks. These outputs were chosen to partially separate the effects of Zn and H₂ which, when combined, influence the occurrence of leaks and ruptures as well as the leak rate as discussed in the previous sections. Figure 166 shows a comparison of the occurrence of circumferential crack estimates and Figure 168 shows a comparison of the mean estimate of maximum circumferential crack inner half-length.

The five replicates were also used to estimate an overall mean value and a 95% prediction interval was created around this mean, as seen in Figure 167 and Figure 169. The interval appears to be sufficiently smooth for both outputs to suggest convergence. The widening of the prediction intervals, however, is significant. In Figure 166, Zn mitigation at 20 years appears to slow the initiation of cracks. Corresponding to this change, the prediction interval in Figure 167 widens, indicating the increased uncertainty with the decrease in events (crack initiations). This pattern can also be seen in Figure 168, where H₂ mitigation at 20 years slows the growth of cracks so the average maximum inner half-length decreases. However, because of the decrease in crack initiations from the Zn mitigation, there are fewer crack growth events to use in estimating crack dimension. Hence, the prediction intervals in Figure 169 also widen at about 20 years and are overall wider with respect to the magnitude of the predictions.

The generally narrow prediction interval and the reflection of expected mitigation behavior in the estimate and interval suggest convergence of the results for occurrence of circumferential crack. The reflection of the expected mitigation behavior in the estimates of maximum inner half-length also support the convergence of those results, though better convergence may be desirable depending on application. Using 1000 epistemic samples with LHS and IS on p2543 produces stable estimates of the mean for most of these outputs, regardless of the random seed used in the simulation. Those outputs that merit increased sample sizes or more investigation of importance sampling are not failing to converge due to the random seed choice.

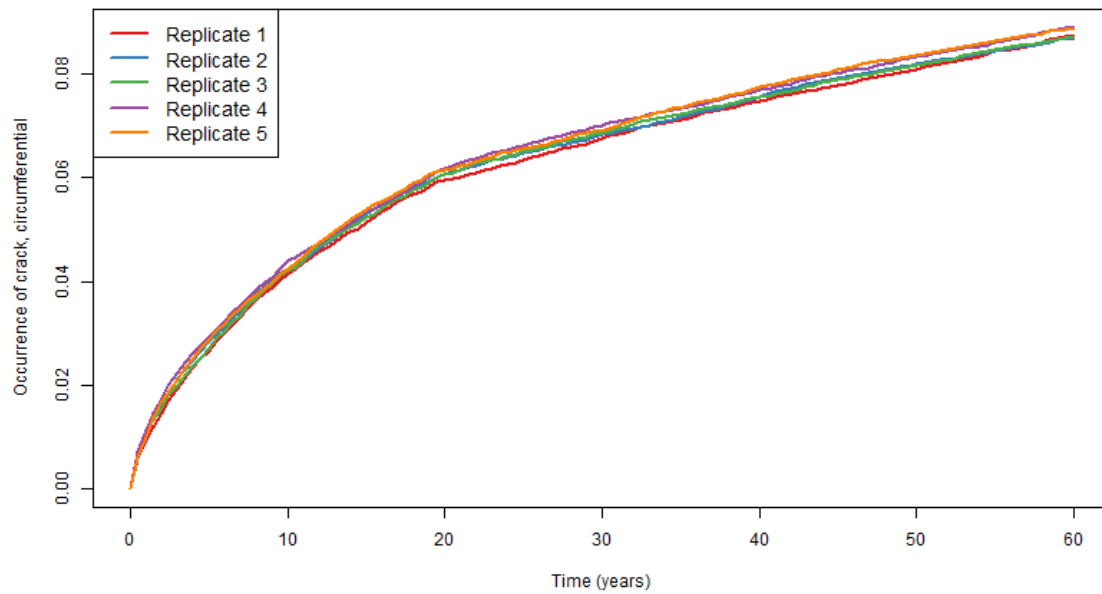


Figure 166: Comparison of estimated mean probability of occurrence of circumferential cracks for five convergence replicates for Scenario 8.

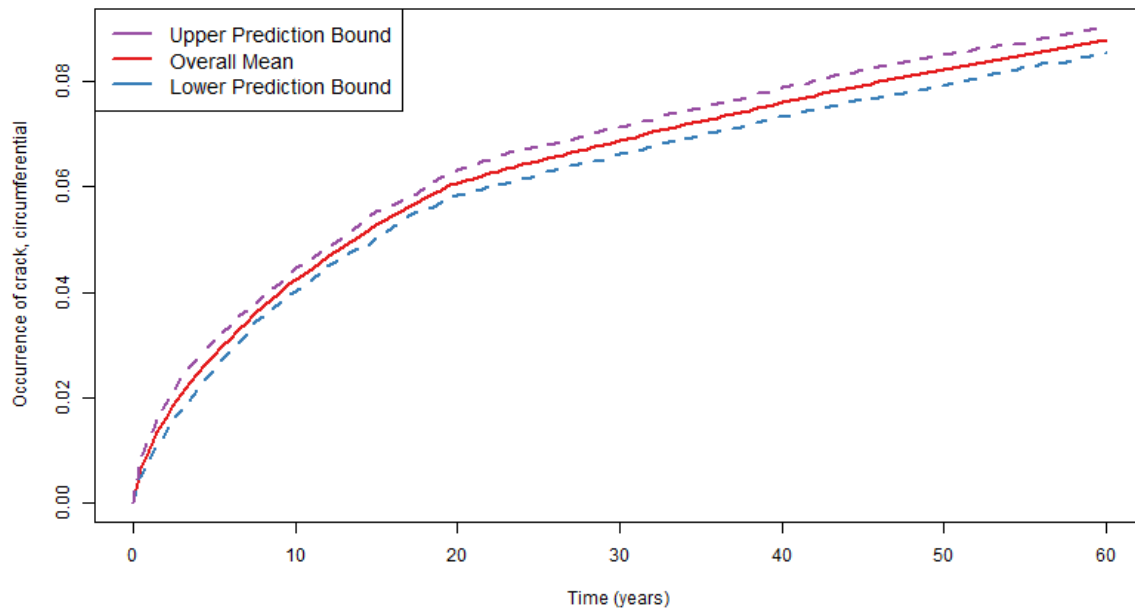


Figure 167: 95% prediction interval over mean probability of circumferential cracks for five replicates for Scenario 8.

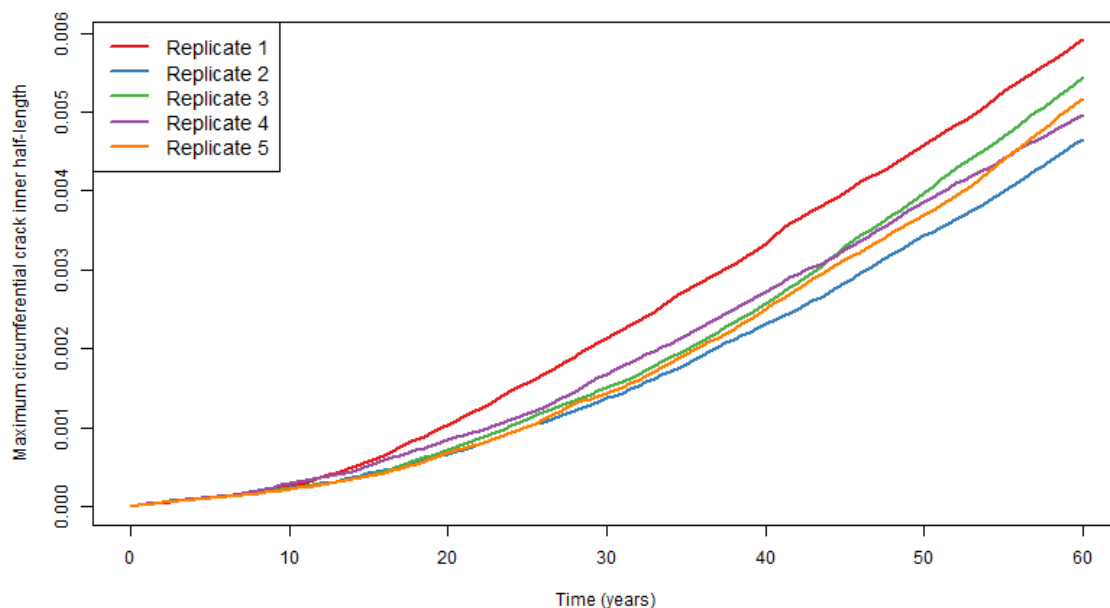


Figure 168: Comparison of the estimated maximum circumferential crack inner half-length for five convergence replicates for Scenario 8.

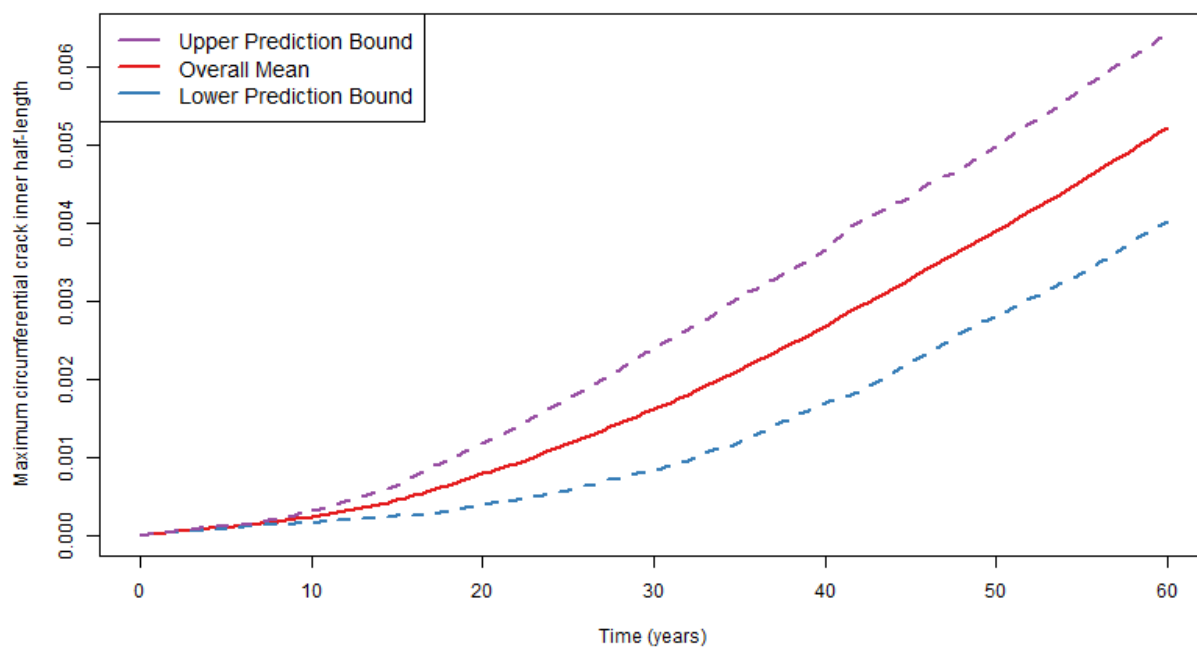


Figure 169: 95% prediction interval over mean estimate of maximum circumferential crack inner half-length for five replicates for Scenario 8.

3.8.6 Comparison to Scenario 3 Results

Scenario 8 and Scenario 3 runs with the same recommended sampling options (LHS IS 1000/50) were compared. The only difference between Scenario 8 Run 4 and Scenario 3 Run 9 is the inclusion of Zn and H₂ mitigation in Scenario 8. Zn mitigation is expected to slow crack initiation, resulting in fewer cracks or the delayed occurrence of cracks. Once cracks initiate, H₂ is expected to slow crack growth, leading to fewer leaks and ruptures (compared to the number of cracks, leaks, and ruptures at the same operating time in the Scenario 3 run). Together, because cracks form later and grow slower, crack dimensions are expected to decrease and the total leak rate is expected to decrease. The figures below demonstrate these results. Crucially, both types of mitigation are applied at 20 years, so the Scenario 3 estimates (blue) and Scenario 8 estimates (red) should agree up to 20 years and then begin to diverge.

As expected, the Scenario 8 estimates of the probability of crack for both axial and circumferential cracks is reduced compared to the estimate from Scenario 3 in Figure 170 and Figure 171. This difference is attributable to the Zn and H₂ mitigation, as the probabilities are identical prior to 20 years. Though no axial ruptures occurred in the Scenario 8 and Scenario 3 simulations, circumferential ruptures are also estimated to be less likely with chemical mitigation (Figure 172). As a consequence both of H₂ slowing crack growth and Zn leading to fewer events to use in estimating crack dimensions, all of the crack dimension estimates in Figure 173 to Figure 175 follow this same pattern. And finally, Figure 176 shows the reduced estimate for the total leak rate after mitigation at 20 years.

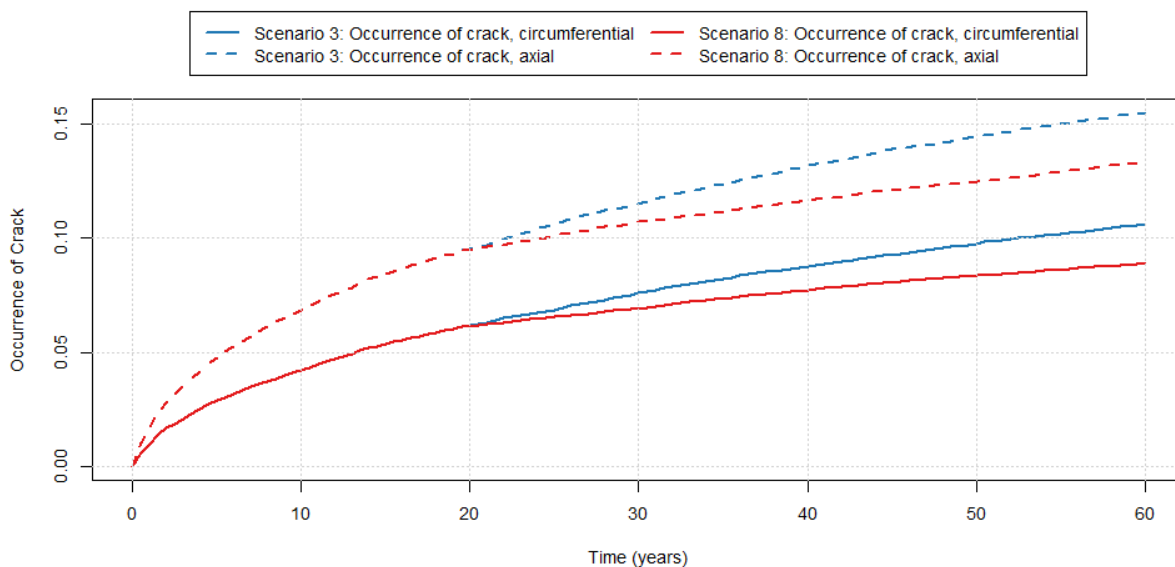


Figure 170: Mean Probability of axial (dashed lines) and circumferential (solid lines) crack occurrence for Scenario 3 (blue lines) and Scenario 8 (red lines).

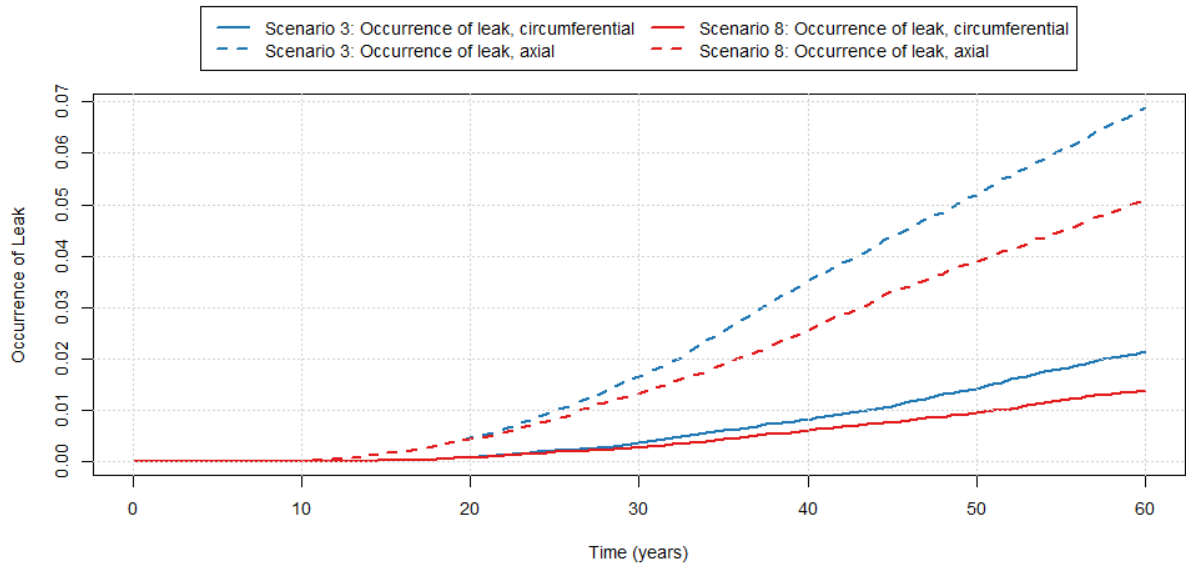


Figure 171: Mean Probability of axial (dashed lines) and circumferential (solid lines) leaks for Scenario 3 (blue lines) and Scenario 8 (red lines).

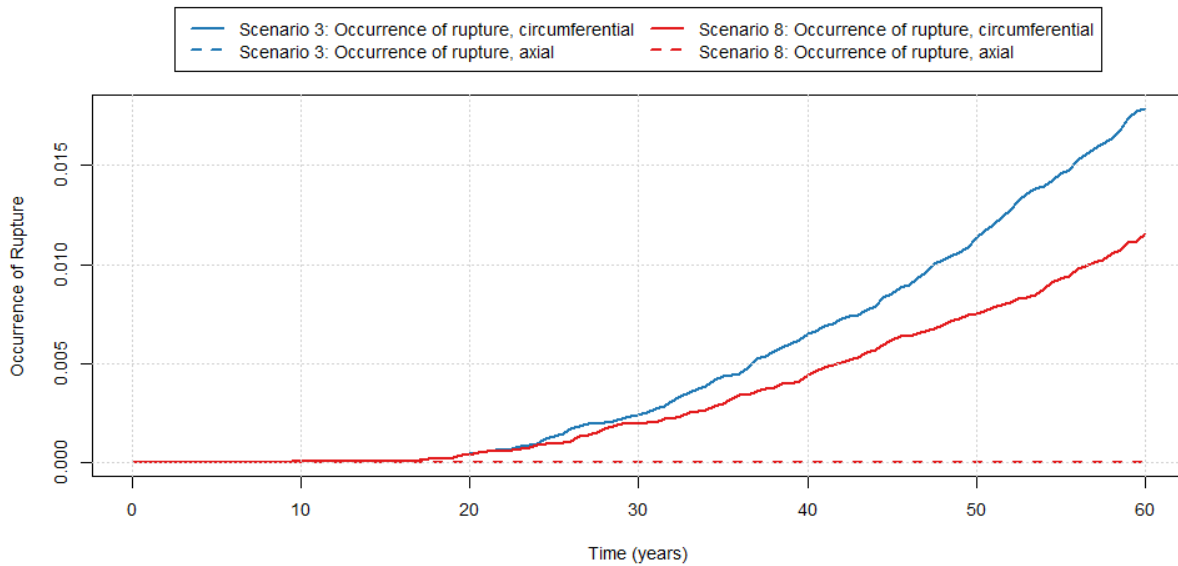


Figure 172: Mean Probability of axial (dashed lines) and circumferential (solid lines) occurrence of rupture for Scenario 3 (blue lines) and Scenario 8 (red lines).

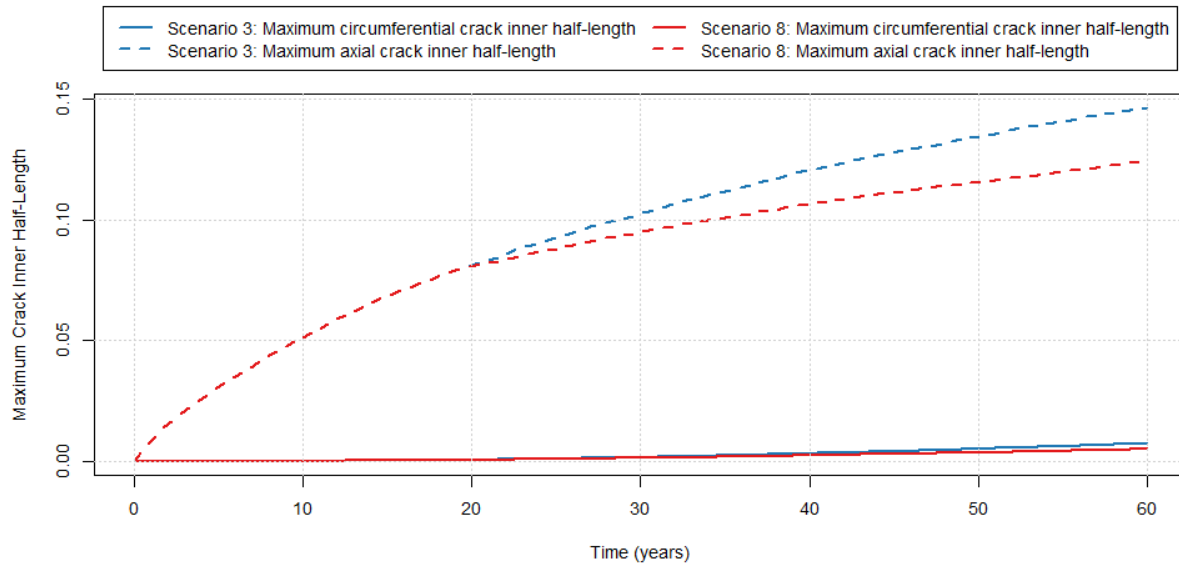


Figure 173: Mean estimates of axial (dashed lines) and circumferential (solid lines) maximum crack inner half-length for Scenario 3 (blue lines) and Scenario 8 (red lines).

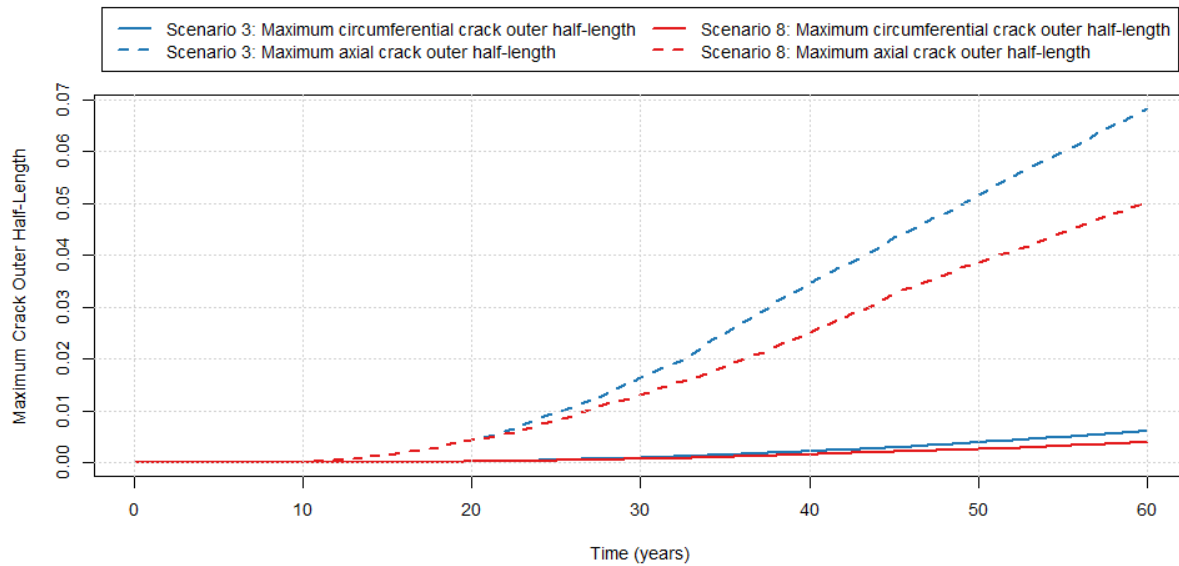


Figure 174: Mean estimates of axial (dashed lines) and circumferential (solid lines) maximum crack outer half-length for Scenario 3 (blue lines) and Scenario 8 (red lines).

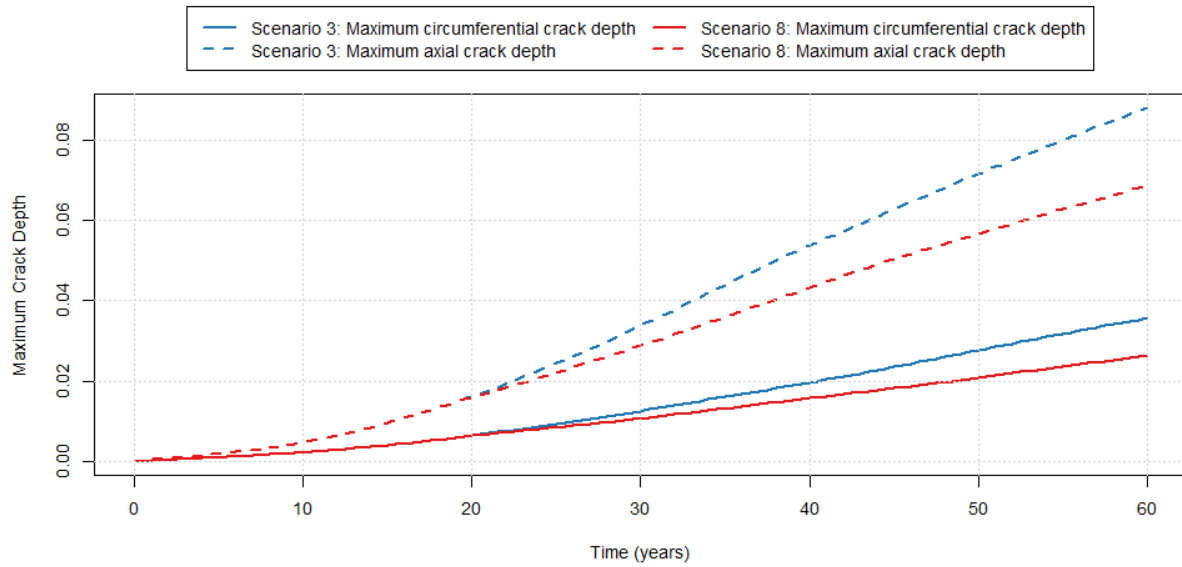


Figure 175: Mean estimates of axial (dashed lines) and circumferential (solid lines) maximum crack depth for Scenario 3 (blue lines) and Scenario 8 (red lines).

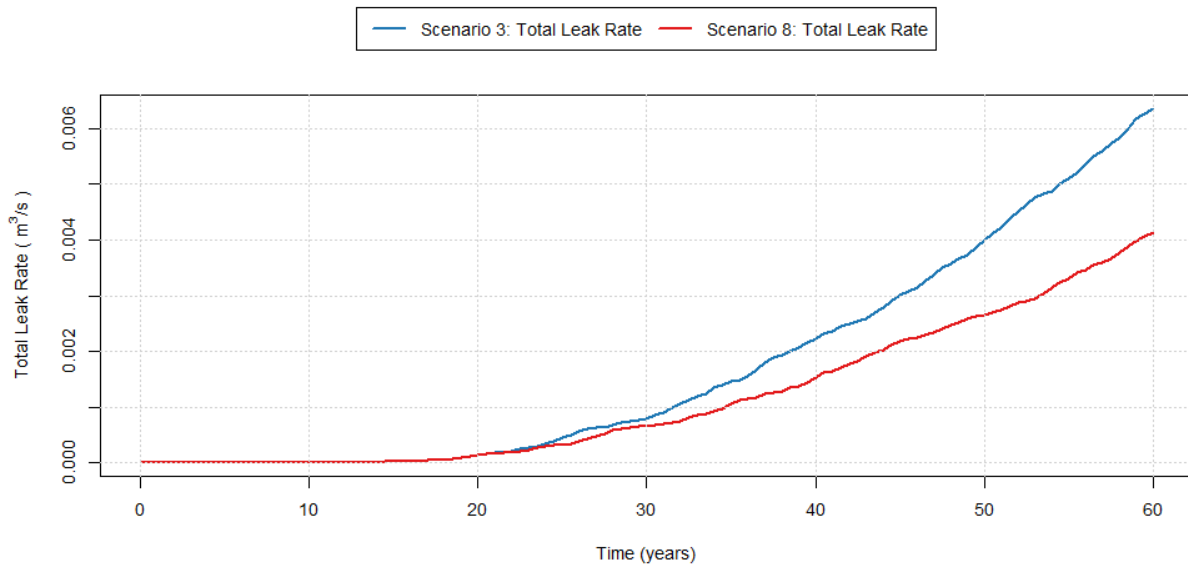


Figure 176: Mean estimates of axial (dashed lines) and circumferential (solid lines) total leak rate for Scenario 3 (blue lines) and Scenario 8 (red lines).

3.9 Scenario 9

3.9.1 Scenario Summary

The results and conclusions of a detailed analysis of sampling schemes and their impact on an understanding of the uncertainty found in the simulation results of Scenario 9 are given in this section.

Scenario 9 uses the following defining options in the xLPR code:

- PWSCC flaw initiation
- Circumferential and axial flaw orientation
- PWSCC flaw growth
- Inlay mitigation at approximately 40 years

The outputs of interest defined for this scenario for both circumferential and axial cracks are:

- Probability of occurrence of crack
- Probability of occurrence of leak
- Probability of occurrence of rupture
- Total leak rate
- Maximum crack depth
- Maximum crack inner half-length
- Maximum crack outer half-length

The comprehensive list of simulations that were used in the analysis of this scenario is shown below. A constant aleatory sample size of 50 was used for each of the sampling schemes considered.

1. 100/50 (epistemic/aleatory sample size) with simple random sampling (SRS)
2. 500/50 with SRS
3. 1000/50 with Latin hypercube sampling (LHS)
4. 1000/50 with LHS and importance sampling (IS) on p2743 (multiplier on DM1 proportionality constant A) with target quantile of 0.95
5. 1000/50 with SRS and IS on p2743 with target quantile of 0.95

The results of the analysis of these sampling schemes are summarized as follows. For Scenario 9, the results of the regression analyses indicate that the multiplier on Direct Model 1 (DM1) proportionality constant A for the PWSCC crack initiation module for mitigation, hereafter referred to as p2743, was the most important input for explaining the variability in the majority of the outputs. Although p2743 was not the most important variable for the common metric used for identify important variables, occurrence of circumferential rupture, the results of the rank regression analysis identified p2743 as the most important variables for the other outputs and was thus used for importance sampling. Overall, results confirm that increasing the epistemic sample size decreases the uncertainty in the mean of the output (e.g., probability of occurrence of axial cracks) as expected.

When inlay mitigation is first implemented at 40 years, occurrences of crack, axial leak, and crack properties are reduced to zero; occurrence of circumferential leak, rupture and the total leak rate

remain fairly constant following mitigation at 40 years. Following inlay mitigation, the rate of increase for axial results becomes greater than Scenario 3 and leads to axial results being greater than those observed for Scenario 3. Scenario 9 circumferential results are less than Scenario 3 results, except for occurrence of circumferential cracks and maximum circumferential crack depth which are greater than Scenario 3 results at year 60.

3.9.2 Analysis Progression

Scenario 3 (Section 3.3) provided a comprehensive progression of the different sampling options available and a thorough analysis from run-to-run. The Scenario 9 analysis progression is not as comprehensive as Scenario 3, but highlights key aspects of the different sampling options to imitate the analysis progression provided in Section 3.3. For a thorough analysis of sampling options, the reader is referred back to Section 3.3. The section for Scenario 9 instead shows the run-to-run progression to demonstrate convergence of the solutions.

Runs 1 and 2 were used to determine if epistemic sample sizes have an impact on result uncertainty. Figure 177 plots the mean probability of occurrence of circumferential rupture for Run 1 and Run 2, along with the 95% confidence intervals (CIs), calculated using the bootstrap method (Section 2.2.2.2), which describes the sampling uncertainty associated with the mean. When inlay mitigation is applied at year 40, the probability remains constant until year 55 where it starts to increase again. The effect of inlay mitigation is discussed further in Section 3.9.6. As seen in Figure 177, the width of the CIs are narrower for the 500/50 SRS run, which is better visualized in Figure 178. From Run 1 to Run 2, the width of the CIs decreases by ~ 0.007 at year 60.

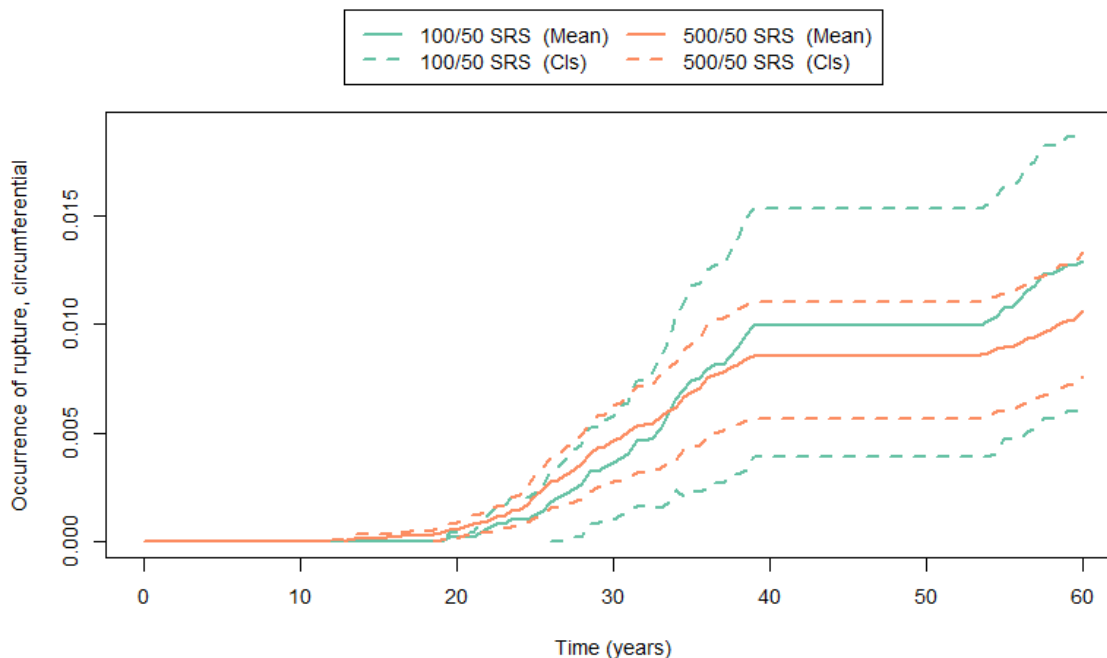


Figure 177: Mean probability of occurrence of circumferential rupture (solid line) and 95% confidence intervals (dashed lines) for Scenario 9, Run 1 (green) and Run 2 (orange).

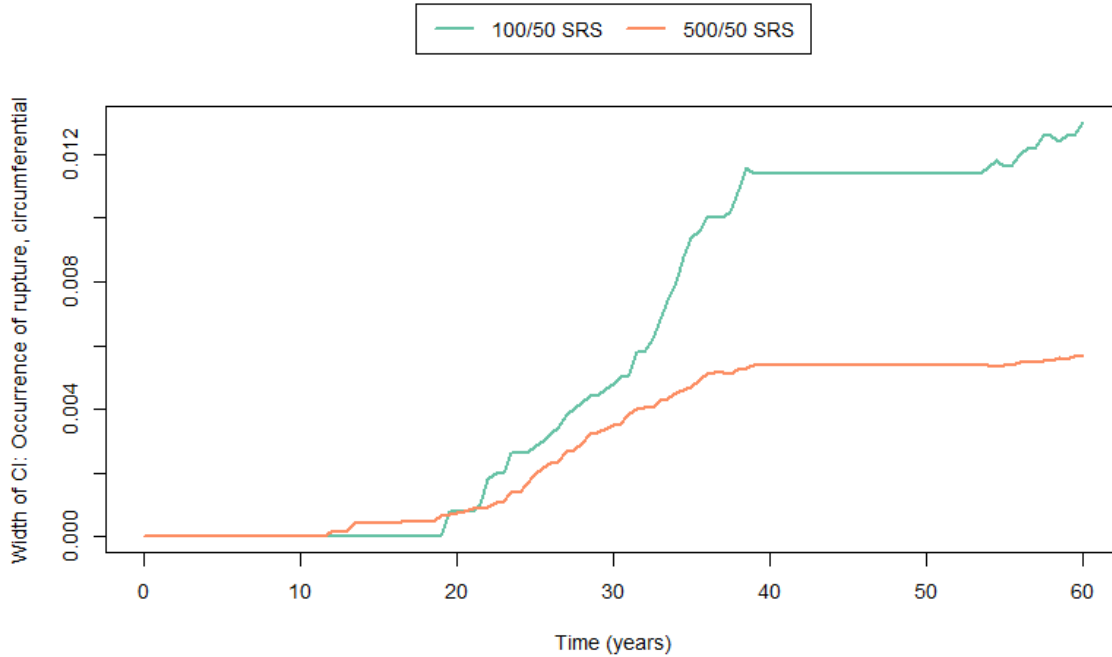


Figure 178: Width of 95% confidence interval for probability of occurrence of circumferential rupture for Scenario 9, Run 1 (green) and Run 2 (orange).

Based on the results seen in Figure 177 and Figure 178, and the analysis progression described in Section 3.3.2, the next progression in the analysis was to increase the epistemic sample size to 1000 and sample using LHS. The justification to move from SRS to LHS is described in Section 4.3. The results for Run 3 are plotted in Figure 179. The LHS confidence bounds presented in Figure 179 are calculated as SRS confidence bounds as described in the methods section and thus cannot serve as direct evidence comparing these two sampling methods (i.e., LHS and SRS).

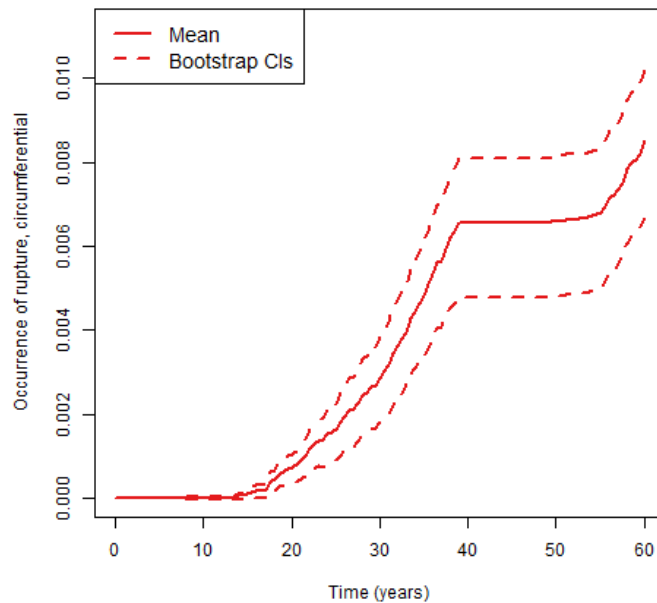


Figure 179: Mean probability of occurrence of circumferential rupture (solid line) and 95% confidence intervals (dashed lines) for Scenario 9, Run 3.

The results of preliminary runs (Runs 1 through 3) identify variable p2743, the multiplier on DM1 proportionality constant A (mitigation), as the most important variable to the majority of the scenario outputs. The scatterplots for these variables show that p2743 has a definitive relationship with occurrence of circumferential cracks as higher values of this variable correlate to higher probabilities of occurrence, as seen in Figure 180. A rank regression analysis was used to confirm the importance of variable p2743 to occurrence of circumferential cracks. The results of this rank regression analysis are shown in Table 10, which compares the standardized rank regression coefficient (SRRC) for the five most significant variables for each result found in Run 3. The probability of occurrence of axial rupture is not shown in this table because there were not any axial ruptures observed for this scenario.

For occurrence of circumferential rupture, p2743 is not the most important variable, as seen in Table 10. The relationship between p2743 and circumferential rupture output is shown in Figure 181. Although the circumferential rupture output is the metric typically used for selection of important variables, Table 10 illustrates that p2743 is the most important variable across the majority of variables and was thus selected instead of p2543.

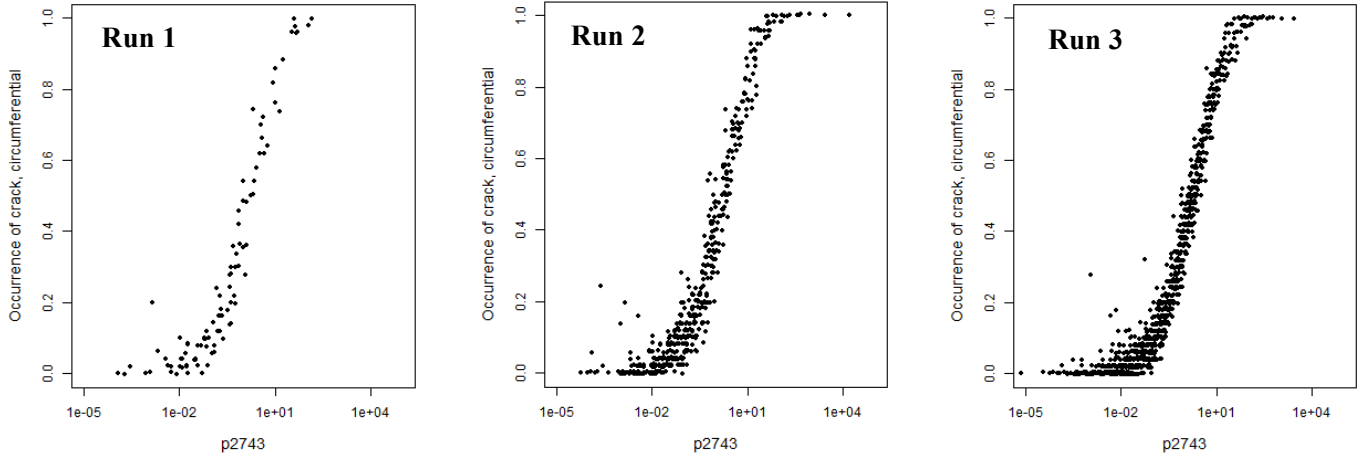


Figure 180: Scatter plots for variable p2743 (multiplier on DM1 proportionality constant A) and the probability of occurrence for circumferential crack for Scenario 9, Run 1 (left), Run 2 (center), and Run 3 (right).

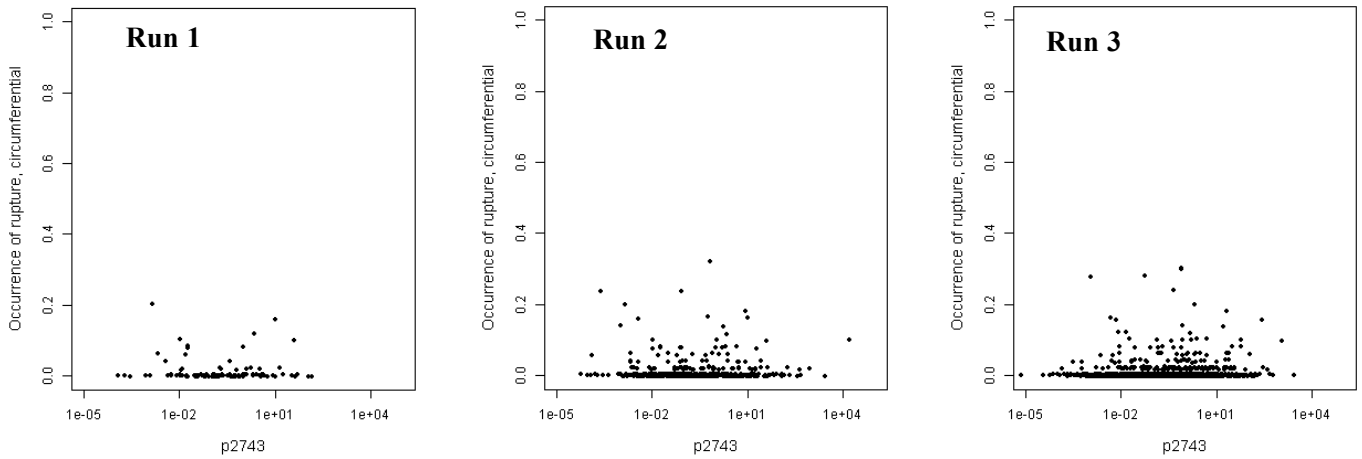


Figure 181: Scatter plots for variable p2743 (multiplier on DM1 proportionality constant A) and the probability of occurrence for circumferential rupture for Scenario 9, Run 1 (left), Run 2 (center), and Run 3 (right).

Table 10: Summary of the stepwise rank regression analysis results for all outputs at year 60 for Scenario 9, Run 3.

Variable Identifier	Variable Name	SRRC	Variable Identifier	Variable Name	SRRC
Probability of Occurrence of Axial Cracks – $R^2 = 0.973$			Maximum Axial Crack Depth – $R^2 = 0.922$		
p2743	Multiplier proport. Const. A (DM1), (Mitigation)	0.986	p2743	Multiplier proport. Const. A (DM1), (Mitigation)	0.933
p4351	Hoop WRS Post-mitigation	0.071	p4351	Hoop WRS Post-mitigation	0.137
p1102	Pipe Wall Thickness	-0.019	p2592	Comp-to-Comp Variab Factor, fcomp	0.118
p2708	Material Init J-Resist Exponent, m, (Mitigation)	0.012	p2595	Charact Width of Peak vs ECP, c	0.115
p2543	Multiplier proport. Const. A (DM1)	0.010	p2792	Comp-to-Comp Variab Factor, fcomp, (Mitigation)	0.115
Probability of Occurrence of Circumferential Cracks – $R^2 = 0.940$			Maximum Circumferential Crack Depth – $R^2 = 0.850$		
p2743	Multiplier proport. Const. A (DM1), (Mitigation)	0.966	p2743	Multiplier proport. Const. A (DM1), (Mitigation)	0.882
p4353	Axial WRS Post-mitigation	0.039	p2792	Comp-to-Comp Variab Factor, fcomp, (Mitigation)	0.137
p5502	Log Reg Slope Param, beta_1 (circ)	-0.039	p2592	Comp-to-Comp Variab Factor, fcomp	0.121
p5107	Depth-sizing Bias Term, a (axial)	-0.038	p1102	Pipe Wall Thickness	-0.105
p1102	Pipe Wall Thickness	-0.038	p2543	Multiplier proport. Const. A (DM1)	0.096
Probability of Axial Leak – $R^2 = 0.801$			Maximum Axial Inner Half-Length – $R^2 = 0.965$		
p2743	Multiplier proport. Const. A (DM1), (Mitigation)	0.813	p2743	Multiplier proport. Const. A (DM1), (Mitigation)	0.979
p2592	Comp-to-Comp Variab Factor, fcomp	0.251	p4351	Hoop WRS Post-mitigation	0.082
p2595	Charact Width of Peak vs ECP, c	0.220	p2792	Comp-to-Comp Variab Factor, fcomp, (Mitigation)	0.043
p4351	Hoop WRS Post-mitigation	0.184	p1102	Pipe Wall Thickness	-0.025
p2594	Peak-to-Valley ECP Ratio - 1, P-1	-0.160	p2543	Multiplier proport. Const. A (DM1)	0.017
Probability of Circumferential Leak – $R^2 = 0.359$			Maximum Circumferential Inner Half-Length – $R^2 = 0.724$		
p2743	Multiplier proport. Const. A (DM1), (Mitigation)	0.302	p2743	Multiplier proport. Const. A (DM1), (Mitigation)	0.776
p2592	Comp-to-Comp Variab Factor, fcomp	0.286	p2543	Multiplier proport. Const. A (DM1)	0.186
p2543	Multiplier proport. Const. A (DM1)	0.259	p2792	Comp-to-Comp Variab Factor, fcomp, (Mitigation)	0.168
p2792	Comp-to-Comp Variab Factor, fcomp, (Mitigation)	0.211	p2592	Comp-to-Comp Variab Factor, fcomp	0.133
p1102	Pipe Wall Thickness	-0.153	p1102	Pipe Wall Thickness	-0.116
Probability of Circumferential Rupture – $R^2 = 0.308$			Maximum Axial Outer Half-Length – $R^2 = 0.801$		
p2543	Multiplier proport. Const. A (DM1)	0.328	p2743	Multiplier proport. Const. A (DM1), (Mitigation)	0.810
p2592	Comp-to-Comp Variab Factor, fcomp	0.288	p2592	Comp-to-Comp Variab Factor, fcomp	0.256
p1102	Pipe Wall Thickness	-0.178	p2595	Charact Width of Peak vs ECP, c	0.225
p2743	Multiplier proport. Const. A (DM1), (Mitigation)	0.130	p4351	Hoop WRS Post-mitigation	0.185
p2594	Peak-to-Valley ECP Ratio - 1, P-1	-0.126	p2594	Peak-to-Valley ECP Ratio - 1, P-1	-0.162
Total Leak Rate – $R^2 = 0.628$			Maximum Circumferential Outer Half-Length – $R^2 = 0.868$		
p2743	Multiplier proport. Const. A (DM1), (Mitigation)	0.603	p2592	Comp-to-Comp Variab Factor, fcomp	0.294
p2592	Comp-to-Comp Variab Factor, fcomp	0.321	p2743	Multiplier proport. Const. A (DM1), (Mitigation)	0.289
p2543	Multiplier proport. Const. A (DM1)	0.237	p2543	Multiplier proport. Const. A (DM1)	0.266
p2595	Charact Width of Peak vs ECP, c	0.208	p2792	Comp-to-Comp Variab Factor, fcomp, (Mitigation)	0.202
p2594	Peak-to-Valley ECP Ratio - 1, P-1	-0.171	p1102	Pipe Wall Thickness	-0.150

Although p2743 is not the most significant variable for occurrence of circumferential rupture, it is still one of the important variables and importance sampling on this variable should lead to increased occurrences of circumferential rupture. High values of p2743 are not well sampled within the input space using traditional SRS or LHS and importance sampling was applied at the 95th quantile for the same reasoning described in Section 3.3.2. This sampling selection led to a larger number of circumferential ruptures, as shown in Figure 182 for Run 3 and Run 4. Figure 182 shows

that differences are not observed until after inlay mitigation is implemented at 40 years. This is expected because p2743 is associated with the inlay rather than the weld (as was the case for Scenario 3) which doesn't affect the results until after inlay mitigation is used.

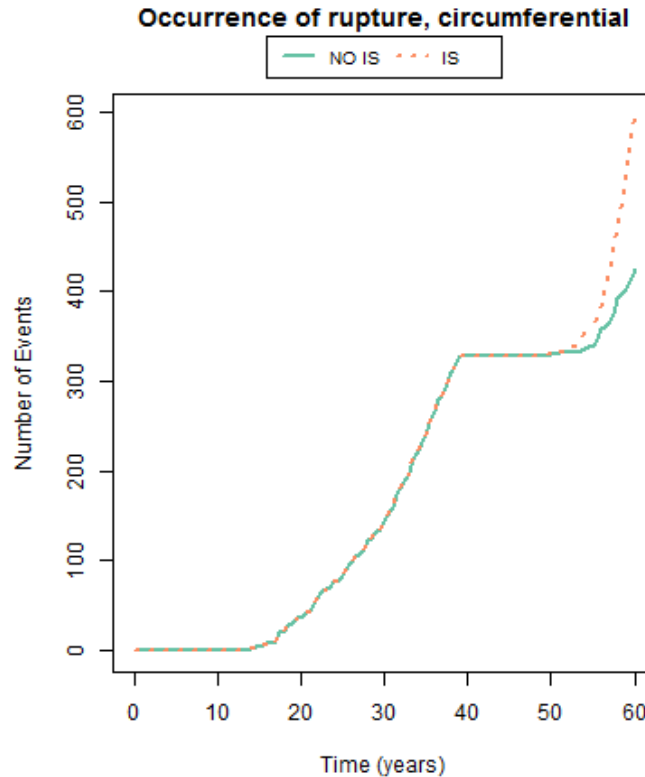


Figure 182: Number of circumferential rupture events occurring for Scenario 9, Run 3 (NO IS) and Run 4 (IS).

Figure 183 and Figure 184 compare Runs 3 and 4 with and without importance sampling for a sample size of 1000/50 using LHS. Effects of importance sampling are not significant with regards to mean probability and the uncertainty associated with the mean probability. A separate study on the importance of the quantile selected for importance sampling, described in Section 4.2, found that 0.95 was not the most efficient selection in terms of decreasing the width of confidence intervals. However, the width of these confidence intervals is narrower in early time when importance sampling is used because the probability of ruptures over this time period is very small. This is the situation in which importance sampling is expected to have the most benefit.

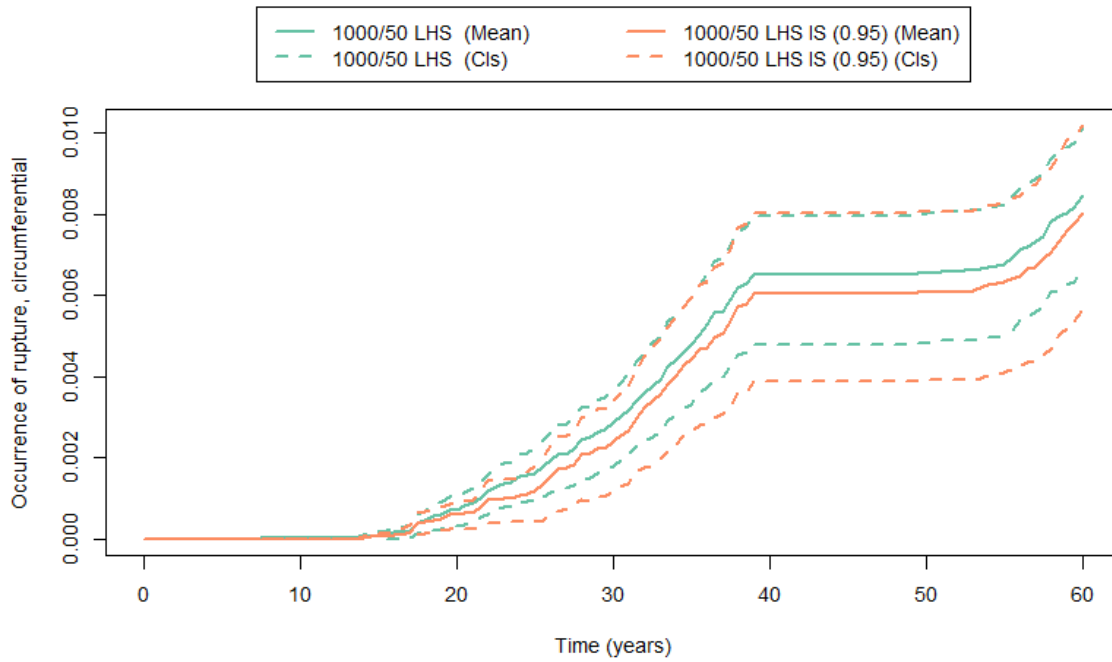


Figure 183: Mean probability of occurrence of circumferential rupture (solid line) and 95% confidence intervals (dashed lines) for Scenario 9, Run 3 (green) and Run 4 (orange).

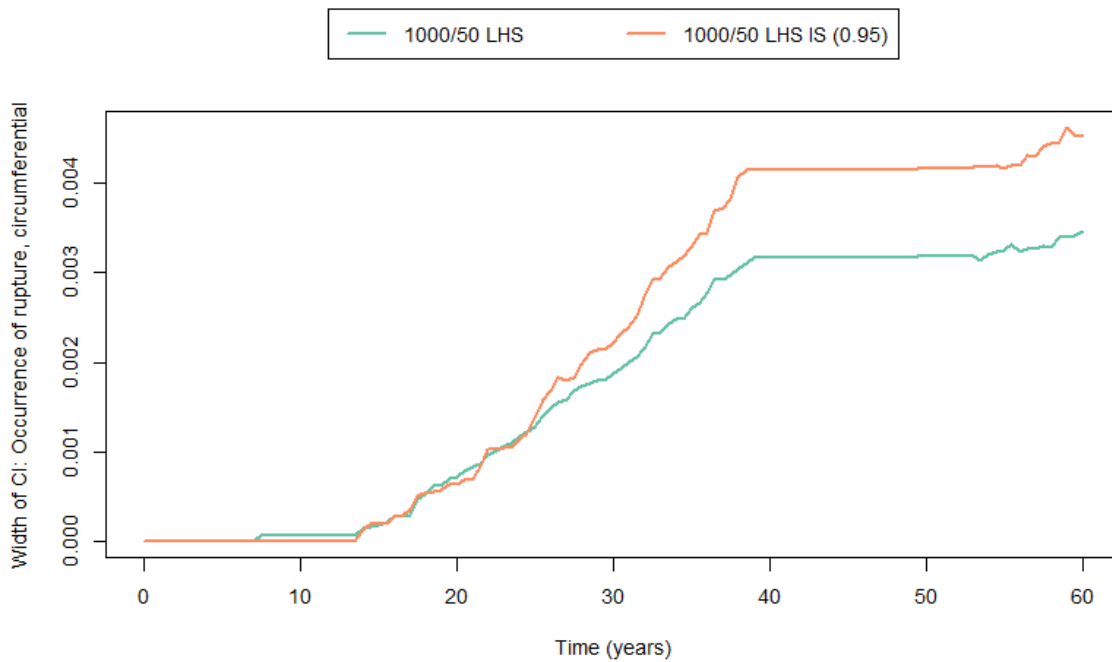


Figure 184: Width of 95% confidence interval for probability of occurrence of circumferential rupture for Scenario 9, Run 3 (green) and Run 4 (orange).

3.9.3 Overall Results

Figure 185 compares the results for the mean probability of occurrence of circumferential ruptures for Runs 1 through 5. All of the runs using SRS, including SRS with IS, diverge from the LHS results for occurrence of circumferential ruptures. Results using LHS are expected to provide better estimates of the mean, as is described in Section 4.3. The divergence of SRS runs without IS can be attributed to the dependence of this sampling method on the selected random seed, as is studied in Section 4.4. However, the use of IS generally removes this dependence. However, the variable used for importance sampling for this scenario, p_{2743} , is not the most important variable for occurrence of circumferential ruptures, meaning that the impact of IS on result stability is diminished.

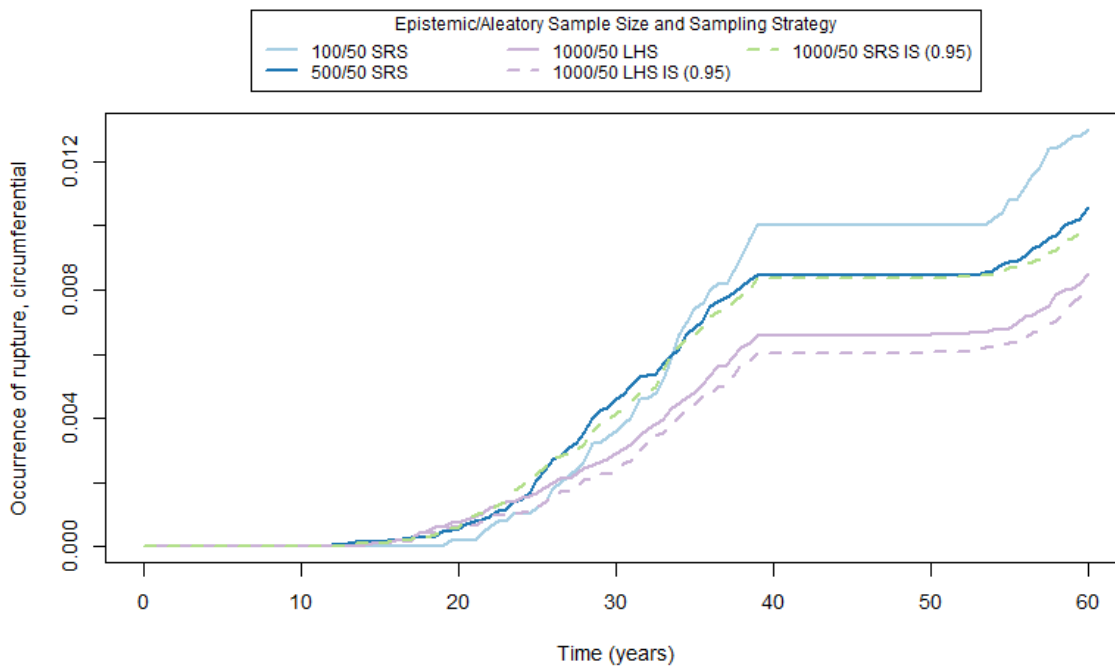


Figure 185: Mean probability of occurrence of circumferential rupture for Scenario 9, Runs 1 through 5.

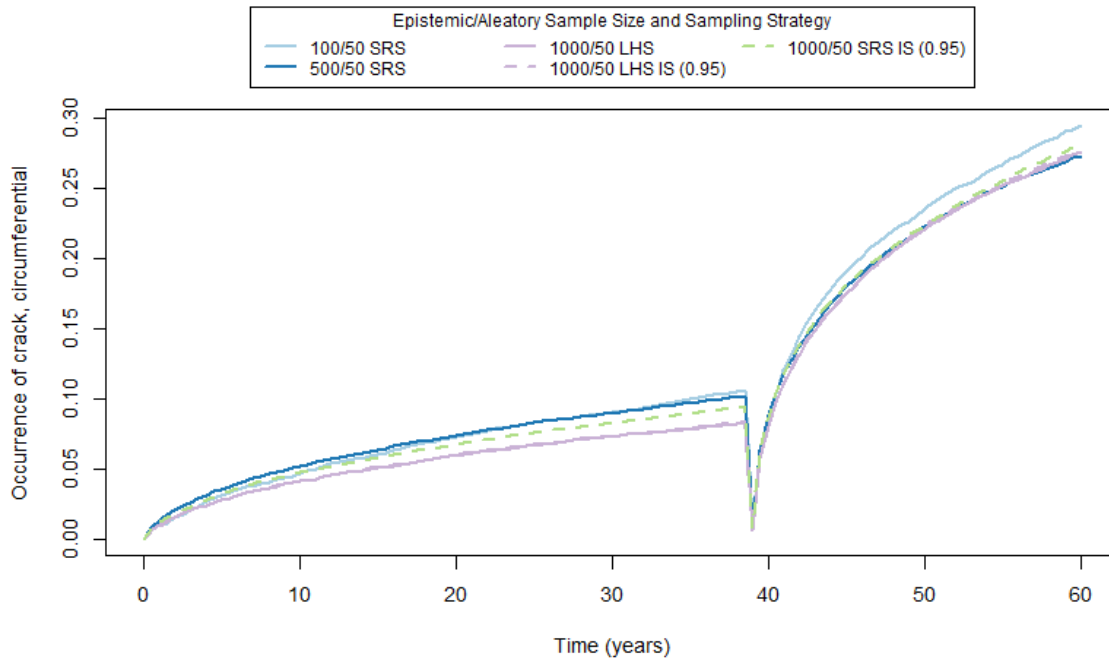


Figure 186: Mean probability of occurrence of circumferential crack for Scenario 9, Runs 1 through 5.

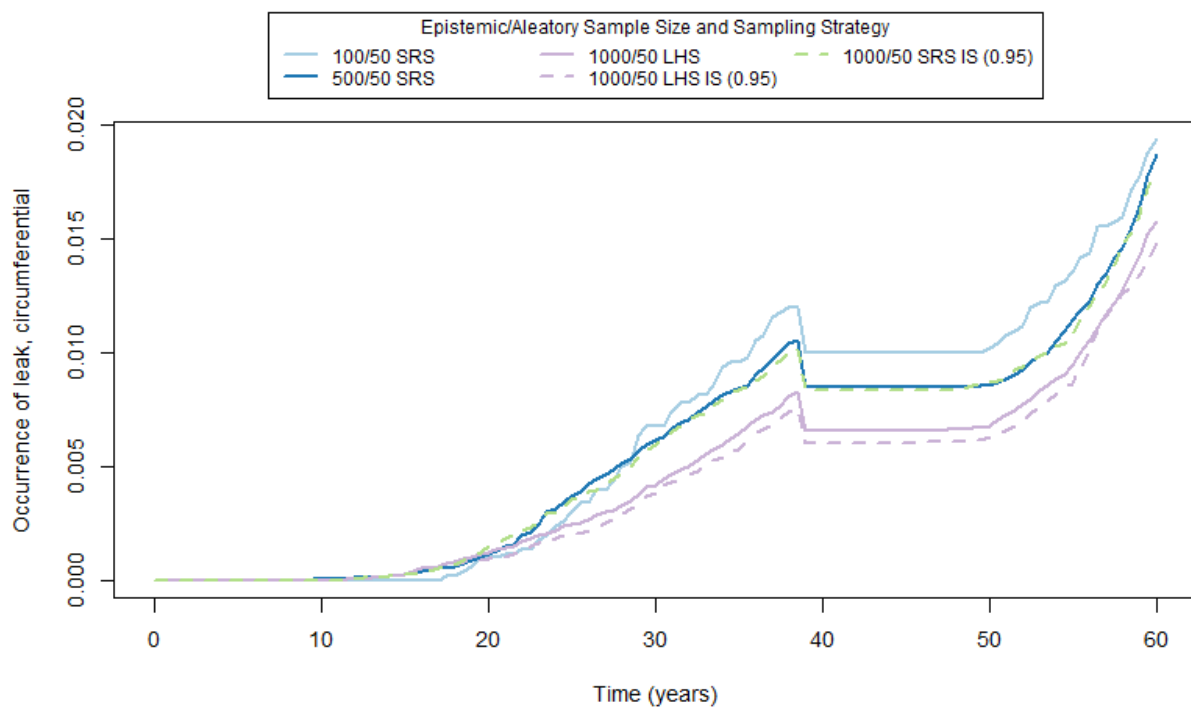


Figure 187: Mean probability of occurrence of circumferential leak for Scenario 9, Runs 1 through 5.

3.9.4 Epistemic Uncertainty Analysis

An uncertainty analysis was conducted for the sampling scheme used in Run 4 (1000 epistemic and 50 aleatory samples and LHS with IS on p2543) with the goal of characterizing epistemic uncertainty in the probabilities of occurrence.

Plots of the estimated probabilities of occurrence of circumferential crack and rupture are displayed in Figure 188 and Figure 189, respectively. These figures show the probability of occurrence for each epistemic realization (grey lines), the mean (red line), and the 5th, 50th, and 95th percentiles (blue, green, and purple lines, respectively). As noted in Section 2.5.4 the percentiles shown here describe the actual variation found in the particular set of epistemic results.

Figure 188 shows that at year 60, 5% of epistemic realizations predict a probability of occurrence of circumferential cracks that is greater than or equal to 0.94. Figure 189 shows that 5% of epistemic realizations have a probability of occurrence of circumferential rupture that is greater than or equal to 0.04 by year 60.

The results for the 95th percentile, along with the mean, for both axial leak and cracks are better visualized in Figure 190. As seen in Figure 190, 5% (95th percentile) of epistemic realizations rapidly approach a probability of occurrence of axial cracks equal to one following inlay mitigation at year 40. By year 45, the 95th percentile is equal to one. As mentioned in previous scenarios, when extreme quantiles are of interest to the user (such as the 5th and 95th percentiles), larger sample sizes may be needed.

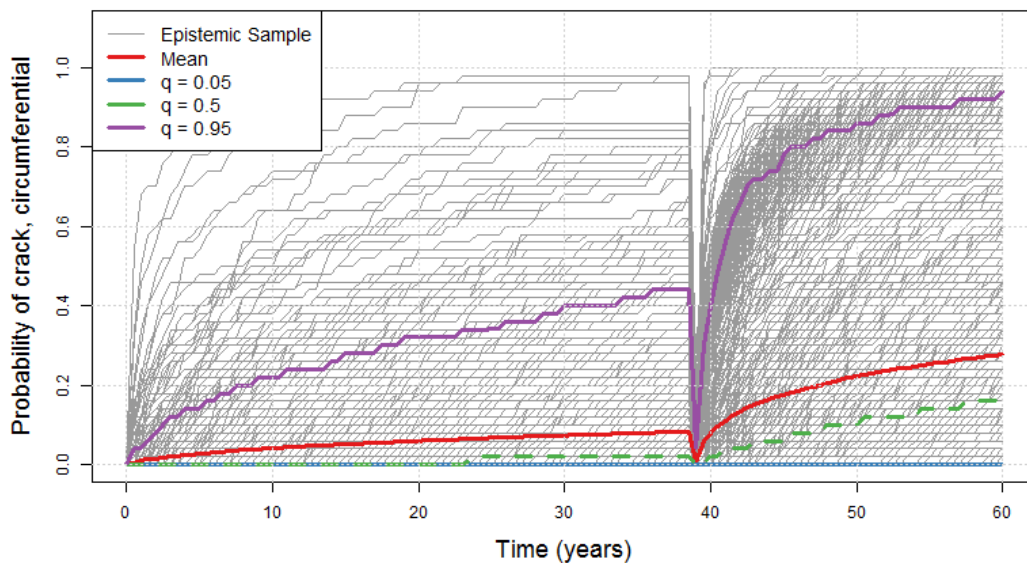


Figure 188: Probability of occurrence of circumferential cracks for each epistemic realization (grey), the mean (red), and the 5th (blue), 50th (green), and 95th (purple) percentiles for Scenario 9, Run 4.

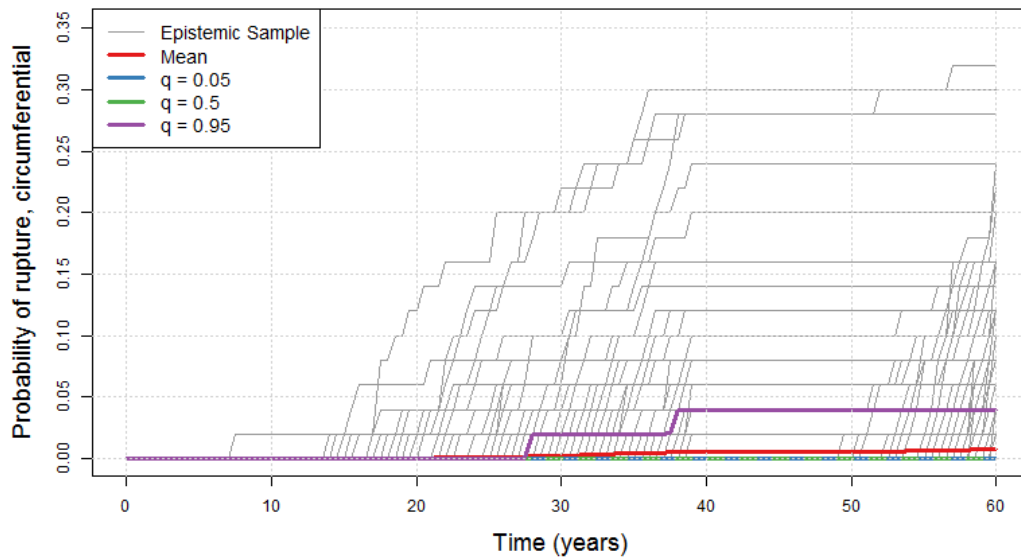


Figure 189: Probability of occurrence of circumferential rupture for each epistemic realization (grey), the mean (red), and the 5th (blue), 50th (green), and 95th (purple) percentiles for Scenario 9, Run 4.

Figure 190 compares the mean and the 95th percentile of the probability of occurrence for circumferential leaks and ruptures. The mean and 95th percentile for both results are approximately equal throughout the simulation time, until about year 55 when the occurrence of leak increases while the occurrence of rupture remains relatively constant. For the 95th percentile, time delay between leaks and ruptures is greater than for the mean.

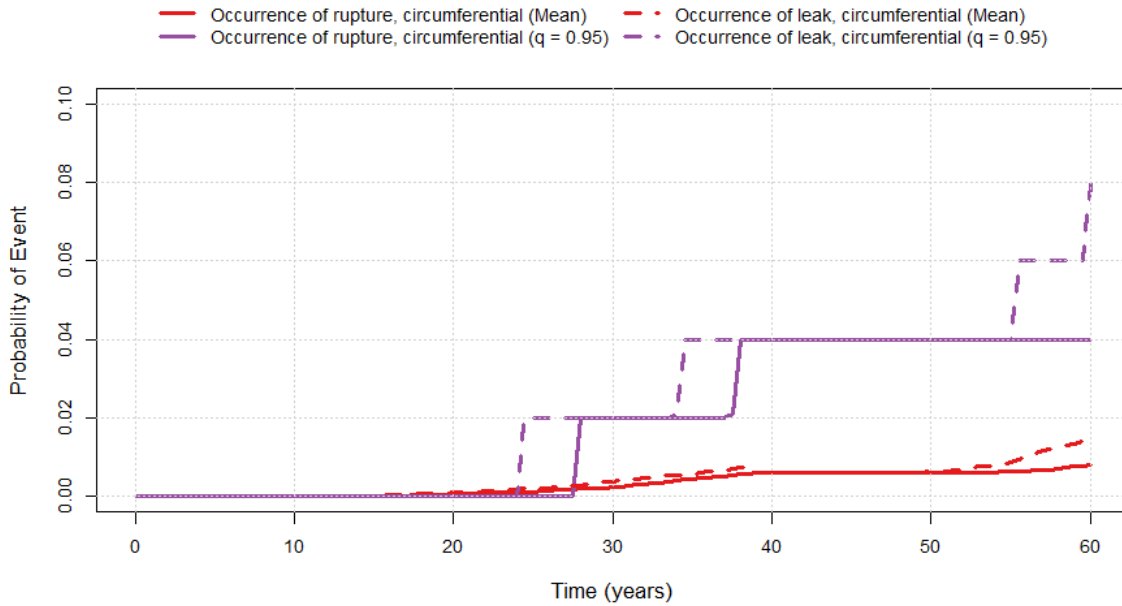


Figure 190: Mean (red) and 95th percentile (purple) for occurrence of circumferential leak (dashed) and rupture (solid) for Scenario 9, Run 4.

3.9.5 Convergence Analysis

The convergence of the sampling scheme for Scenario 4, Run 4 (1000 epistemic samples with LHS and IS) was analyzed using five replicate simulations with different random number seeds for both the epistemic and aleatory loops. Figure 191 shows the mean probability of occurrence of circumferential cracks for each replicate and Figure 192 shows the mean probability of occurrence of circumferential rupture for each replicate. A 95% prediction interval was created around the overall mean of the five replicates, as seen in Figure 193 and Figure 194 for occurrence of circumferential crack and rupture, respectively.

The width of the 95% prediction interval is relatively narrow over the course of the simulation time for occurrence of circumferential cracks, particularly after inlay mitigation at 40 years (see Figure 193). The width of the 95% prediction interval is about 0.7% at year 60 for occurrence of circumferential rupture, as seen in Figure 194.

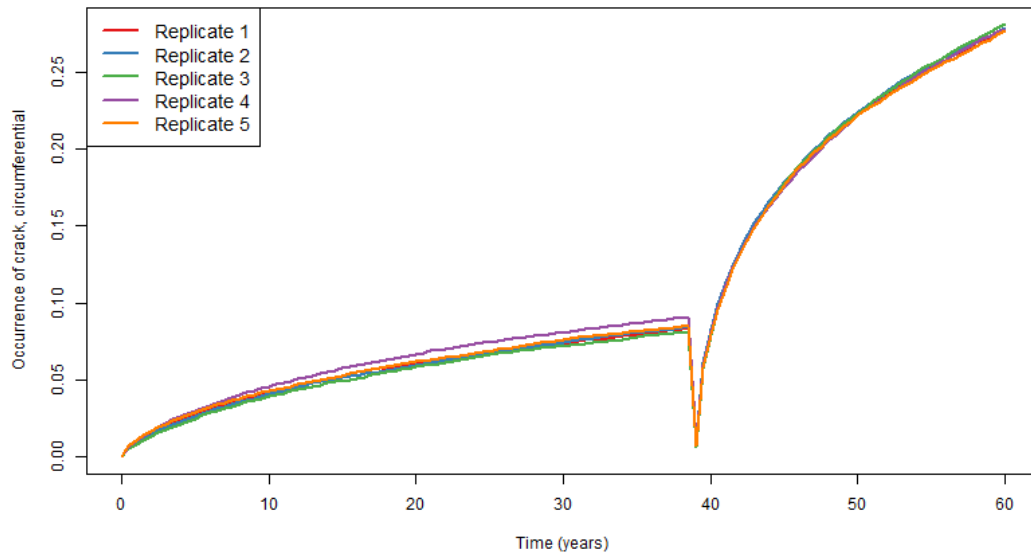


Figure 191: Comparison of estimated mean probability of occurrence of circumferential cracks for five convergence replicates for Scenario 9.

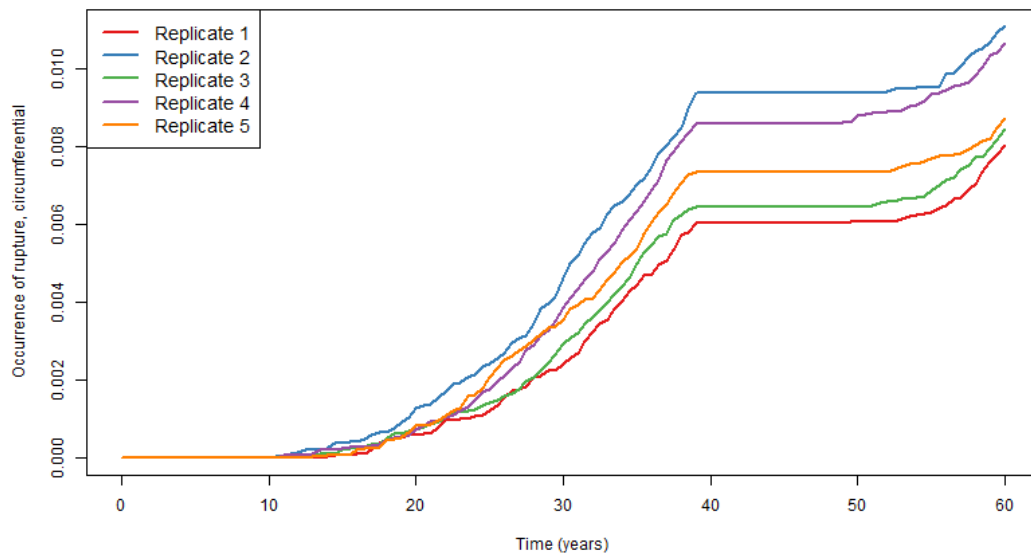


Figure 192: Comparison of estimated mean probability of occurrence of circumferential rupture for five convergence replicates for Scenario 9.

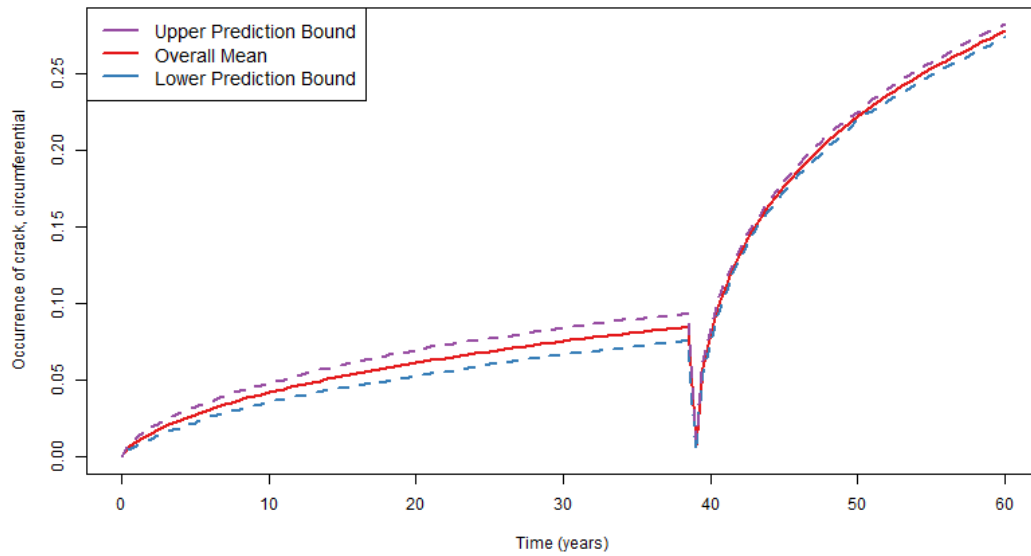


Figure 193: 95% prediction interval over mean probability of circumferential cracks for five replicates for Scenario 9.

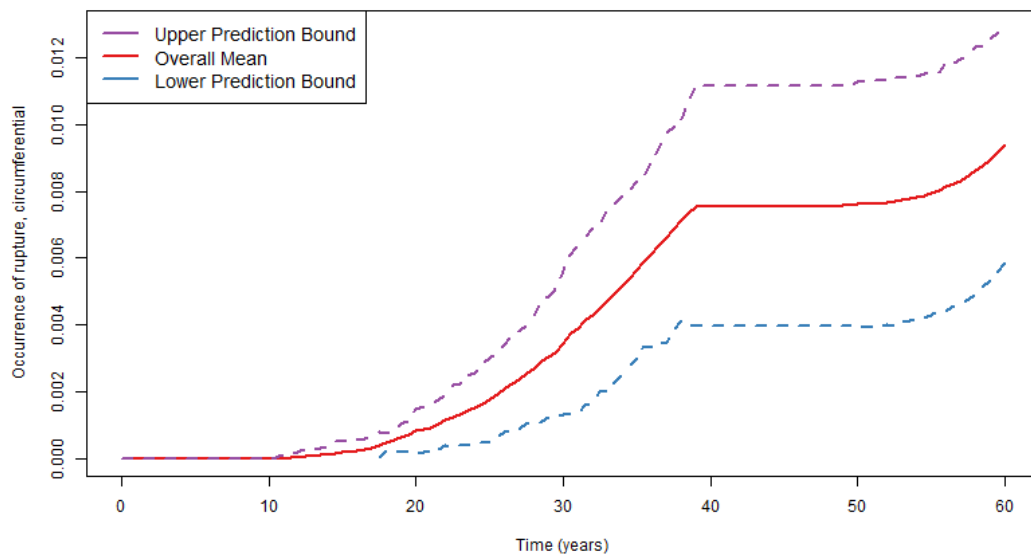


Figure 194: 95% prediction interval over mean probability of circumferential rupture for five replicates for Scenario 9.

3.9.6 Comparison to Scenario 3 Results

Scenario 9 runs were compared to Scenario 3 runs with corresponding sampling selections. To properly compare the two scenarios, Run 9 from Scenario 3 and Run 4 from Scenario 9 were used for the comparison as both runs represent a run with 1000 epistemic samples with LHS and IS.

The effects of inlay mitigation at 40 years are most apparent when compared to Scenario 3, as seen in Figure 195 through Figure 201. Compared to Scenario 3, inlay mitigation results in higher probabilities of occurrence of axial cracks, circumferential cracks, and axial leak at year 60. Closer inspection of the input set suggests that the cause of this increase is due to the inner diameter post-mitigation WRS that results from inlay mitigation. The resulting WRSs are tensile (both hoop and axial WRSs) whereas prior to the mitigation, the WRSs were compressive. This variable contributes to the overall occurrence of cracks and axial leak being greater than Scenario 3, but it is not identified as an important variable because of the relatively small standard deviation associated with its distribution. However, the resulting WRSs do not cause occurrences of circumferential leak or rupture to be greater than Scenario 3.

Crack results for maximum crack inner half-length (iHL), maximum outer half-length (oHL), and maximum crack depth are shown in Figure 198, Figure 199, and Figure 200, respectively. Inlay mitigation causes axial crack behavior to be markedly different for these crack properties, while circumferential crack behavior is approximately the same as Scenario 3.

The total leak rate for Scenario 3 and Scenario 9 are compared in Figure 201. When inlay mitigation is applied at 40 years (Scenario 9), the total leak rate at 60 years is about three times less than Scenario 3 which has no mitigation options.

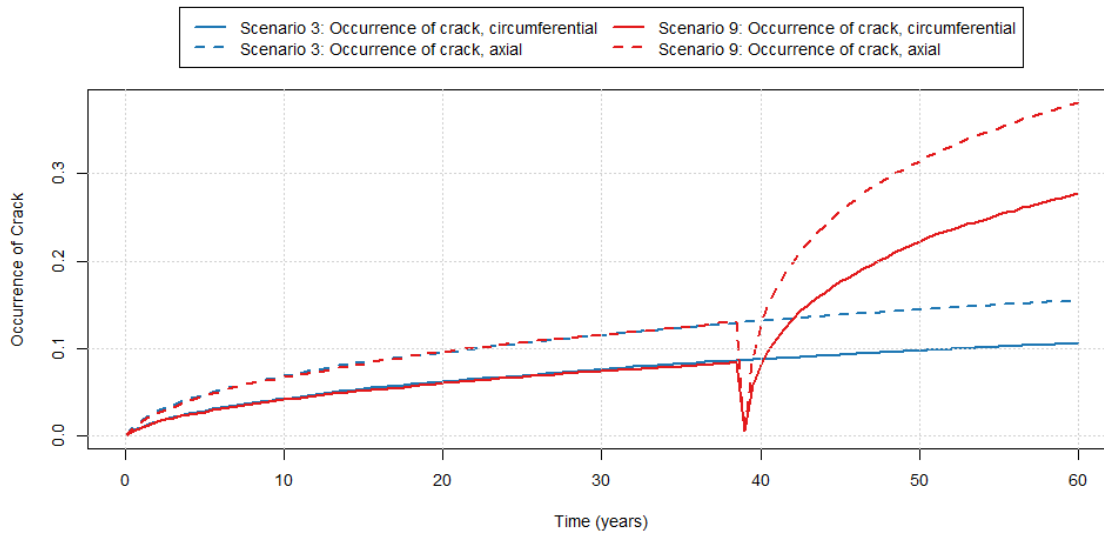


Figure 195: Mean probability of axial (dashed lines) and circumferential (solid lines) crack occurrence for Scenario 3 (blue lines) and Scenario 9 (red lines).

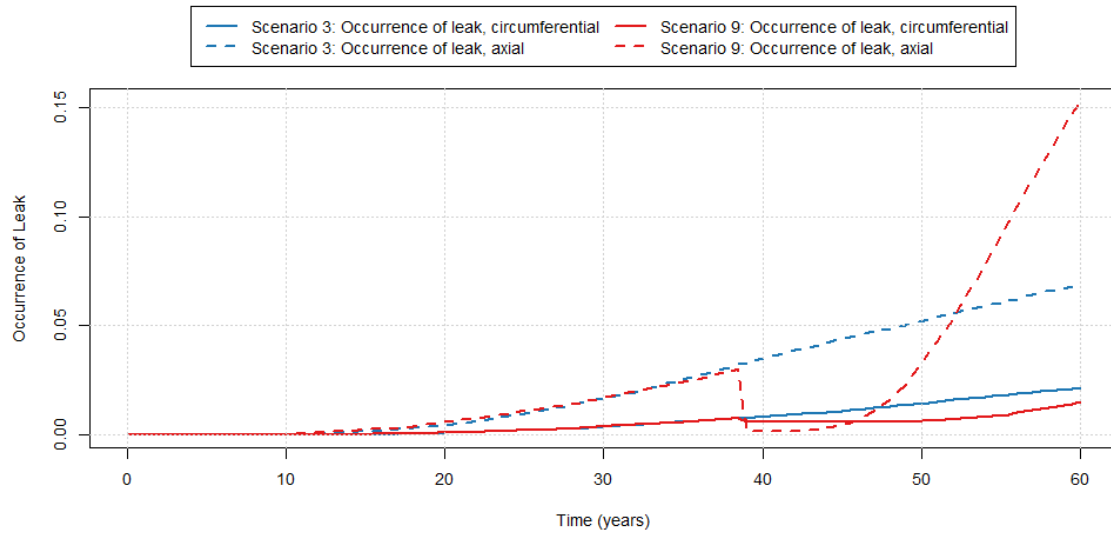


Figure 196: Mean probability of axial (dashed lines) and circumferential (solid lines) leak occurrence for Scenario 3 (blue lines) and Scenario 9 (red lines).

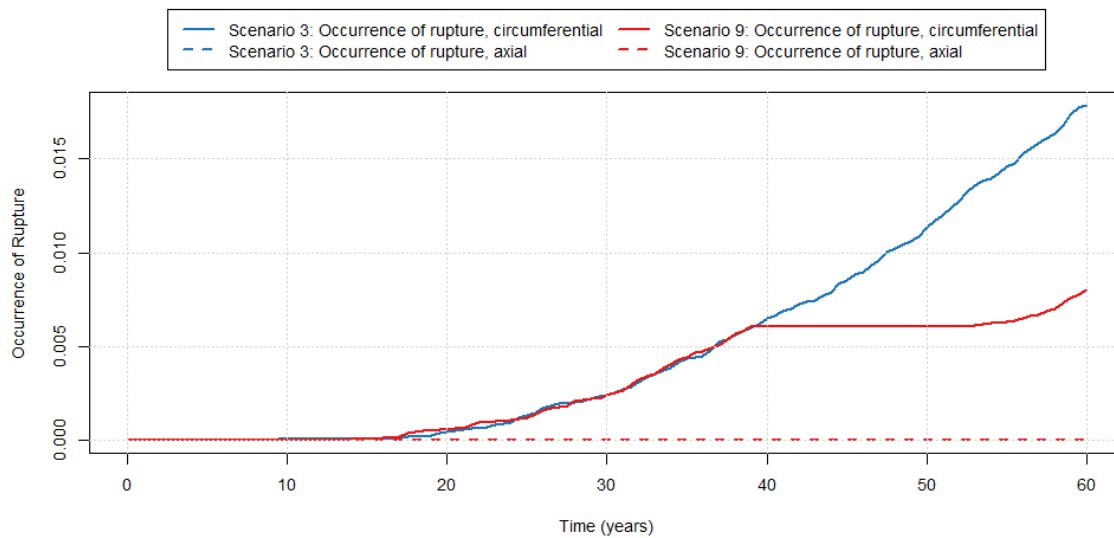


Figure 197: Mean probability of axial (dashed lines) and circumferential (solid lines) rupture occurrence for Scenario 3 (blue lines) and Scenario 9 (red lines).

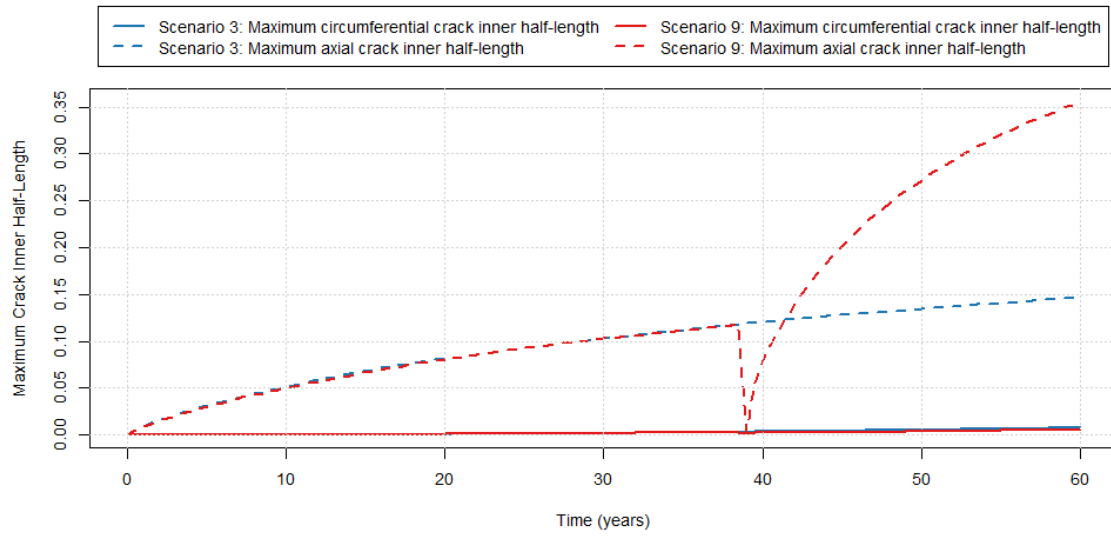


Figure 198: Maximum axial (dashed lines) and circumferential (solid lines) iHL for Scenario 3 (blue lines) and Scenario 9 (red lines).

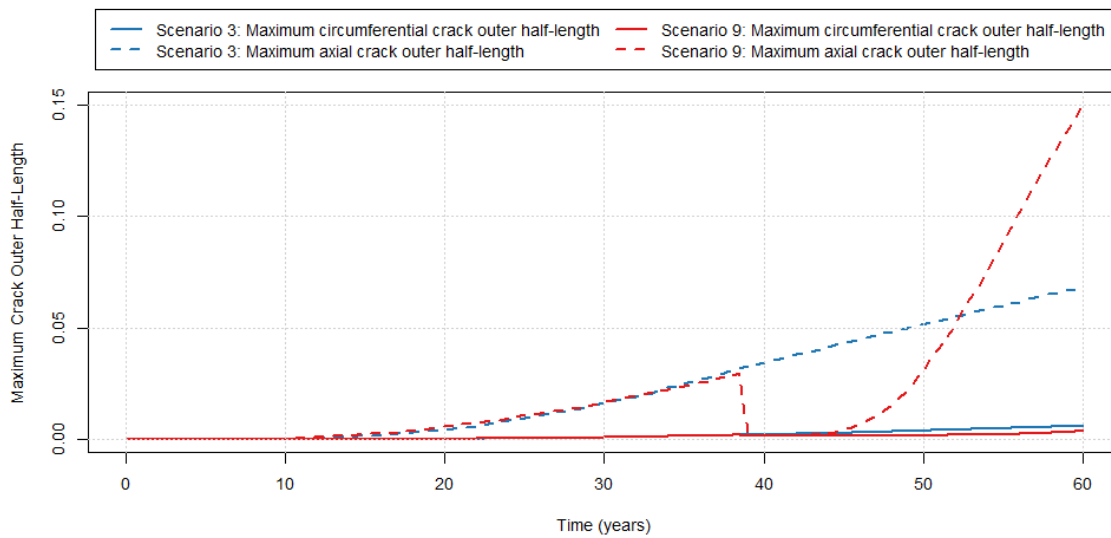


Figure 199: Maximum axial (dashed lines) and circumferential (solid lines) oHL for Scenario 3 (blue lines) and Scenario 9 (red lines).

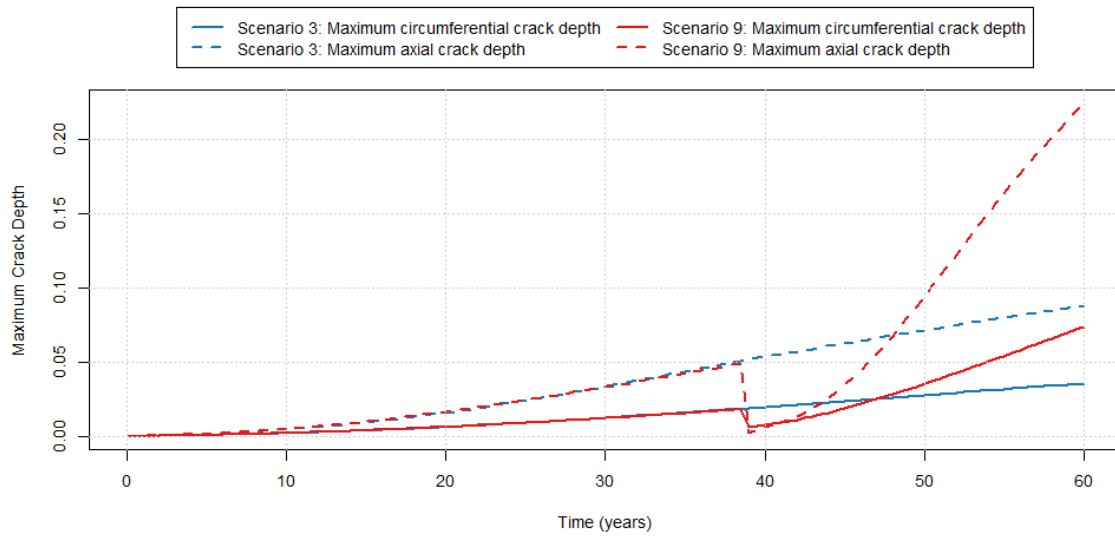


Figure 200: Maximum axial (dashed lines) and circumferential (solid lines) crack depth for Scenario 3 (blue lines) and Scenario 9 (red lines).

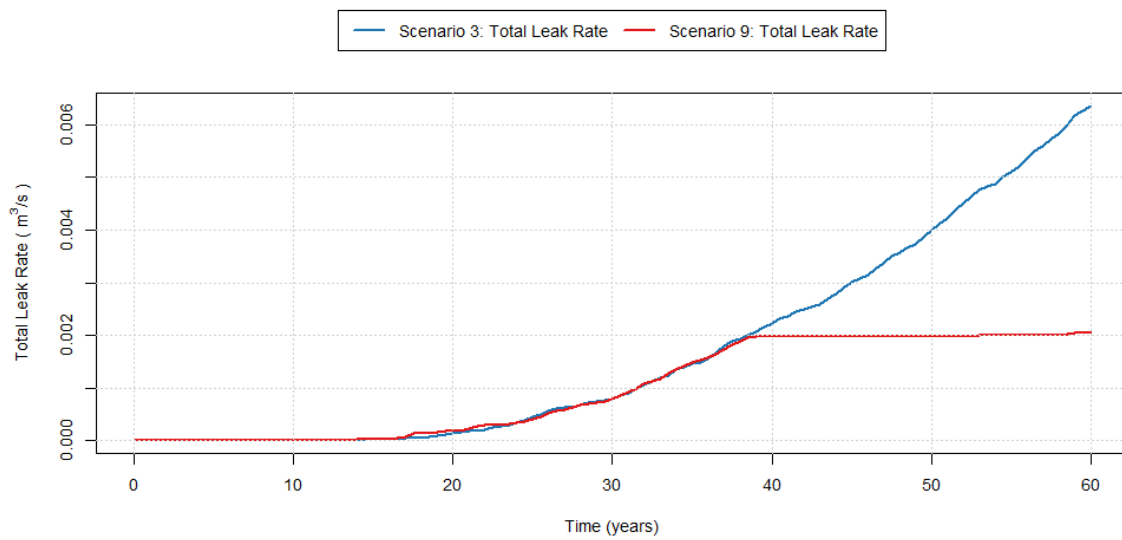


Figure 201: Total leak rate for Scenario 3 (blue line) and Scenario 9 (red line).

3.10 Scenario 10

3.10.1 Scenario Summary

Scenario 10 examines PWSCC and fatigue initiation and growth under full chemical and stress mitigation at 20 years. The results and conclusions of an analysis of sampling schemes and their impact on the overall understanding of the uncertainty found in the simulation results of Scenario 10 are given in this section.

Scenario 10 uses the following defining options in the xLPR code:

- PWSCC and fatigue flaw initiation
- Circumferential and axial flaw orientation
- PWSCC and fatigue flaw growth
- Chemical mitigation with dissolved hydrogen (H₂) and Zinc (Zn) at 20 years
- Stress mitigation using mechanical stress improvement process (MSIP) at approximately 20 years

These code options make Scenario 10 comparable to Scenario 3 which serves as the basis for Scenario 10 but does not include any mitigation. The mitigation options selected for Scenario 10 combine the mitigation strategies for Scenario 4 (MSIP stress mitigation) and Scenario 8 (H₂ and Zn chemical mitigation). These comparisons will be made in Section 3.10.6. Scenario 10 also includes fatigue initiation and growth. As was found in Scenario 1 (fatigue growth) and Scenario 11 (fatigue initiation and growth), this mechanism does not strongly contribute to crack initiation and growth when compared to PWSCC, especially given the transients defined for Scenario 10. More extreme transient definitions were employed for Scenario 11 (see Section 3.11) in order to produce fatigue initiations.

The outputs of interest defined for this scenario for both circumferential and axial cracks are:

- Probability of occurrence of crack
- Probability of occurrence of leak
- Probability of occurrence of rupture
- Total leak rate
- Maximum crack depth
- Maximum crack inner half-length
- Maximum crack outer half-length

A comprehensive list of the simulations that were used in the analysis of this scenario is shown below. A constant aleatory sample size of 50 realizations was used for each of the sampling schemes considered for this scenario.

1. 100/50 (epistemic/aleatory sample size) with simple random sampling (SRS)
2. 500/50 with SRS
3. 1000/50 with Latin hypercube sampling (LHS)
4. 1000/50 with LHS and importance sampling (IS) on parameter p2543 (multiplier on DM1 proportionality constant A) with target quantile of 0.95
5. 1000/50 with SRS and IS on parameter p2543 with target quantile of 0.95

The selection of variable p2543 (multiplier on PWSCC crack initiation DM1 proportionality constant A) for importance sampling will be motivated by the discussion in Section 3.10.2.

The figures of merit shown in this section display results for the probability of occurrence of circumferential crack ruptures. This result is of interest in the xLPR application space and is representative of the results of the analysis of all of the outputs of interest studied for this scenario.

3.10.2 Analysis Progression

This section describes the run-to-run progression of sampling option selections made for the analysis of Scenario 10. The first three runs of Scenario 10 were simulated to determine an appropriate sample size for the analysis of this scenario. Subsequent runs tested the impact of IS and SRS vs. LHS sampling options.

The comparison of Runs 1 and 2 shows the marked impact of epistemic sample size on result uncertainty.

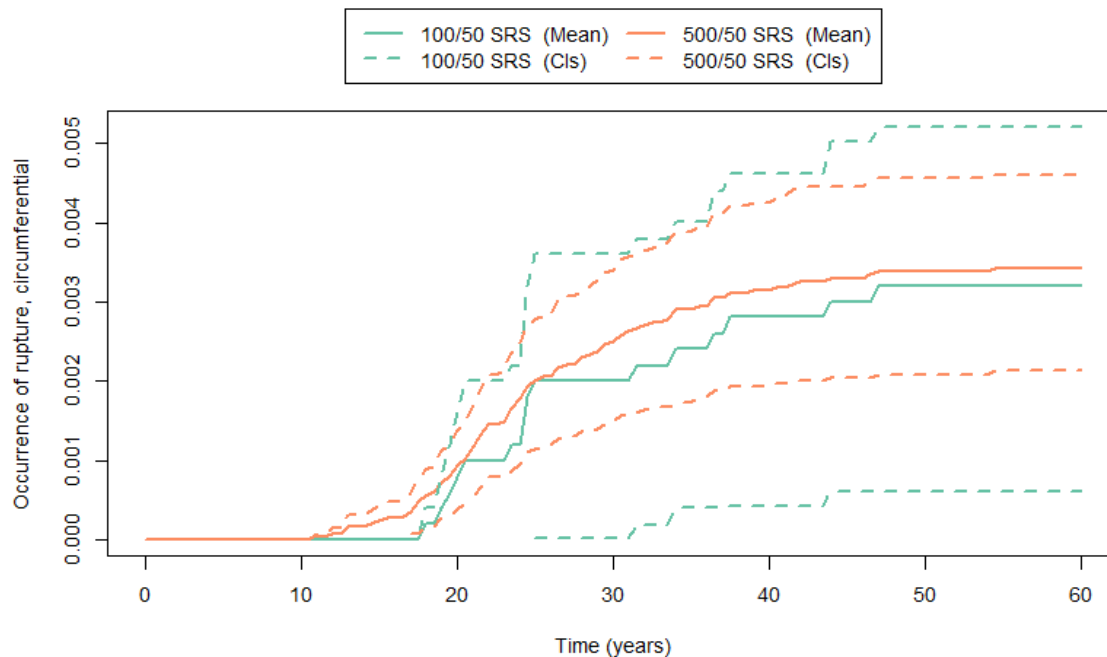


Figure 202: Mean probability of occurrence of circumferential ruptures (solid line) and 95% confidence intervals (dashed lines) for Scenario 10, Run 1 (green) and Run 2 (orange).

A larger sample size smooths the mean result and also decreases the width of the 95% confidence intervals, thus decreasing the uncertainty in the mean result. The use of LHS also decreases the uncertainty in the mean result. This is supported by the case study presented in Section 4.3. The LHS confidence bounds presented in the figure below are calculated as SRS confidence bounds as described in the methods section and thus cannot serve as direct evidence comparing these two sampling algorithms.

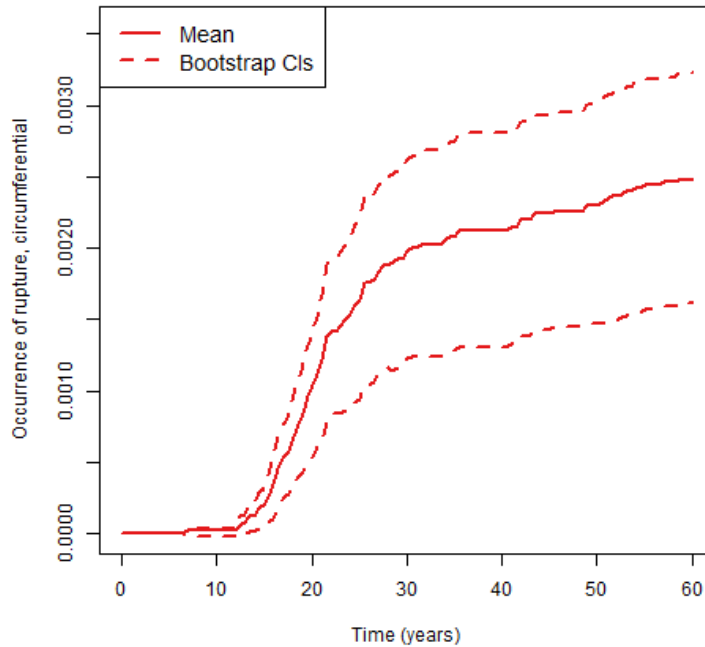


Figure 203: Mean probability of occurrence of circumferential ruptures (solid line) and 95% confidence intervals (dashed lines) for Scenario 10, Run 3.

The analysis of Scenario 3 showed that SRS runs are more susceptible to the selected random seed (Section 4.4). Because Run 3 utilizes LHS, we can be more confident that this is closer to the expected result. The width of the CIs shown in Figure 203 show that the use of a larger sample size in Run 3 coupled with the application of LHS reduces the uncertainty in the mean result for this scenario.

A sensitivity analysis was run on the outputs of Run 3 to determine if any variables could be selected for importance sampling based on how their uncertainty impacts the uncertainty in the results of interest. The results of the stepwise rank regression analysis used to make this determination are shown in Table 11 which compares the standardized rank regression coefficient (SRRC) for the five most significant variables for each result found in Run 3. These results show that p2543 (multiplier on DM1 proportionality constant A) is the most important parameter across all of the outputs of interest for Scenario 10. Subsequent important variables changed based on output and do not have as high of a correlation to the output results based on this regression model. An inspection of the scatterplots of all uncertain variables compared to the outputs of interest supports these results. An example of this is shown for occurrence of circumferential ruptures in Figure 204.

Table 11: Summary of the stepwise rank regression analysis results for all outputs at year 60 for Scenario 10, Run 3.

Variable Identifier	Variable Name	SRRC	Variable Identifier	Variable Name	SRRC
Probability of Occurrence of Axial Cracks – $R^2 = 0.810$			Maximum Axial Crack Depth – $R^2 = 0.774$		
p2543	Multiplier proport. Const. A (DM1)	0.884	p2543	Multiplier proport. Const. A (DM1)	0.838
p4350	Hoop WRS Pre-mitigation	-0.183	p4350	Hoop WRS Pre-mitigation	-0.233
p5106	b (circ)	0.096	p5106	b (circ)	0.091
p5105	a (circ)	0.067	p2594	Peak-to-Valley ECP Ratio - 1, P-1	-0.068
p9002	Surface Crack Dist Rule Modifier	-0.045	p5105	a (circ)	0.067
Probability of Occurrence of Circumferential Cracks – $R^2 = 0.763$			Maximum Circumferential Crack Depth – $R^2 = 0.746$		
p2543	Multiplier proport. Const. A (DM1)	0.862	p2543	Multiplier proport. Const. A (DM1)	0.839
p4352	Axial WRS Pre-mitigation	-0.083	p4352	Axial WRS Pre-mitigation	-0.102
p5401	Intercept, B0 (circ)	0.076	p5401	Intercept, B0 (circ)	0.083
p5402	Slope, B1 (circ)	0.050	p1102	Pipe Wall Thickness	-0.063
p9001	Fatigue Growth CKTH	0.033	p5402	Slope, B1 (circ)	0.059
Probability of Axial Leak – $R^2 = 0.251$			Maximum Axial Inner Half-Length – $R^2 = 0.798$		
p2543	Multiplier proport. Const. A (DM1)	0.360	p2543	Multiplier proport. Const. A (DM1)	0.869
p2592	Comp-to-Comp Variab Factor, fcomp	0.228	p4350	Hoop WRS Pre-mitigation	-0.221
p2594	Peak-to-Valley ECP Ratio - 1, P-1	-0.183	p5106	b (circ)	0.094
p2595	Charact Width of Peak vs ECP, c	0.170	p5105	a (circ)	0.067
p5401	Intercept, B0 (circ)	0.128	p2591	Activation Energy, Qg	0.048
Probability of Circumferential Leak – $R^2 = 0.223$			Maximum Circumferential Inner Half-Length – $R^2 = 0.741$		
p2543	Multiplier proport. Const. A (DM1)	0.323	p2543	Multiplier proport. Const. A (DM1)	0.837
p1102	Pipe Wall Thickness	-0.191	p4352	Axial WRS Pre-mitigation	-0.109
p2592	Comp-to-Comp Variab Factor, fcomp	0.167	p5401	Intercept, B0 (circ)	0.087
p4352	Axial WRS Pre-mitigation	-0.131	p5402	Slope, B1 (circ)	0.061
p2594	Peak-to-Valley ECP Ratio - 1, P-1	-0.114	p2594	Peak-to-Valley ECP Ratio - 1, P-1	-0.059
Probability of Circumferential Rupture – $R^2 = 0.223$			Maximum Axial Outer Half-Length – $R^2 = 0.251$		
p2543	Multiplier proport. Const. A (DM1)	0.323	p2543	Multiplier proport. Const. A (DM1)	0.359
p1102	Pipe Wall Thickness	-0.191	p2592	Comp-to-Comp Variab Factor, fcomp	0.229
p2592	Comp-to-Comp Variab Factor, fcomp	0.167	p2594	Peak-to-Valley ECP Ratio - 1, P-1	-0.185
p4352	Axial WRS Pre-mitigation	-0.131	p2595	Charact Width of Peak vs ECP, c	0.171
p2594	Peak-to-Valley ECP Ratio - 1, P-1	-0.114	p4350	Hoop WRS Pre-mitigation	-0.126
Total Leak Rate – $R^2 = 0.314$			Maximum Circumferential Outer Half-Length – $R^2 = 0.223$		
p2543	Multiplier proport. Const. A (DM1)	0.405	p2543	Multiplier proport. Const. A (DM1)	0.323
p2592	Comp-to-Comp Variab Factor, fcomp	0.245	p1102	Pipe Wall Thickness	-0.189
p2594	Peak-to-Valley ECP Ratio - 1, P-1	-0.165	p2592	Comp-to-Comp Variab Factor, fcomp	0.167
p2595	Charact Width of Peak vs ECP, c	0.161	p4352	Axial WRS Pre-mitigation	-0.132
p1102	Pipe Wall Thickness	-0.129	p2594	Peak-to-Valley ECP Ratio - 1, P-1	-0.114

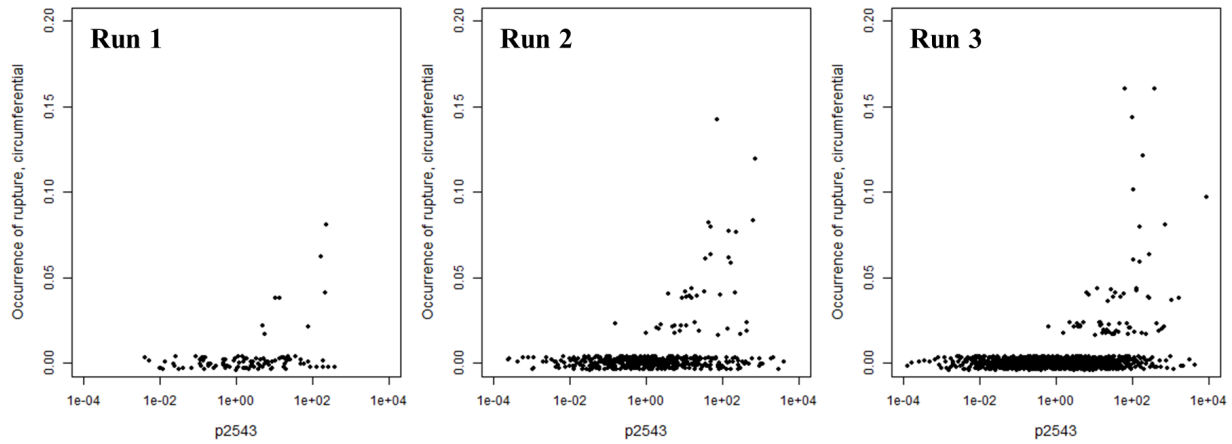


Figure 204: Scatter plots for variable p2543 (multiplier on DM1 proportionality constant A) (plotted on a log scale) and the probability of occurrence for circumferential rupture for Scenario 10, Run 1 (left), Run 2 (center) and Run 3 (right).

Run 4 utilizes importance sampling on variable p2543. Figure 205 and Figure 206 show the impact of this selection on the uncertainty associated with the mean occurrence of circumferential ruptures. The width of the confidence interval does not decrease at 60 years using importance sampling. A separate study on the importance of the quantile selected for importance sampling, described in Section 4.2, found that 0.95 was not the most efficient selection in terms of decreasing the width of confidence intervals. However, the width of these confidence intervals is narrower in early time when importance sampling is used because the probability of ruptures over this time period is very small. This is the situation in which importance sampling is expected to have the most benefit.

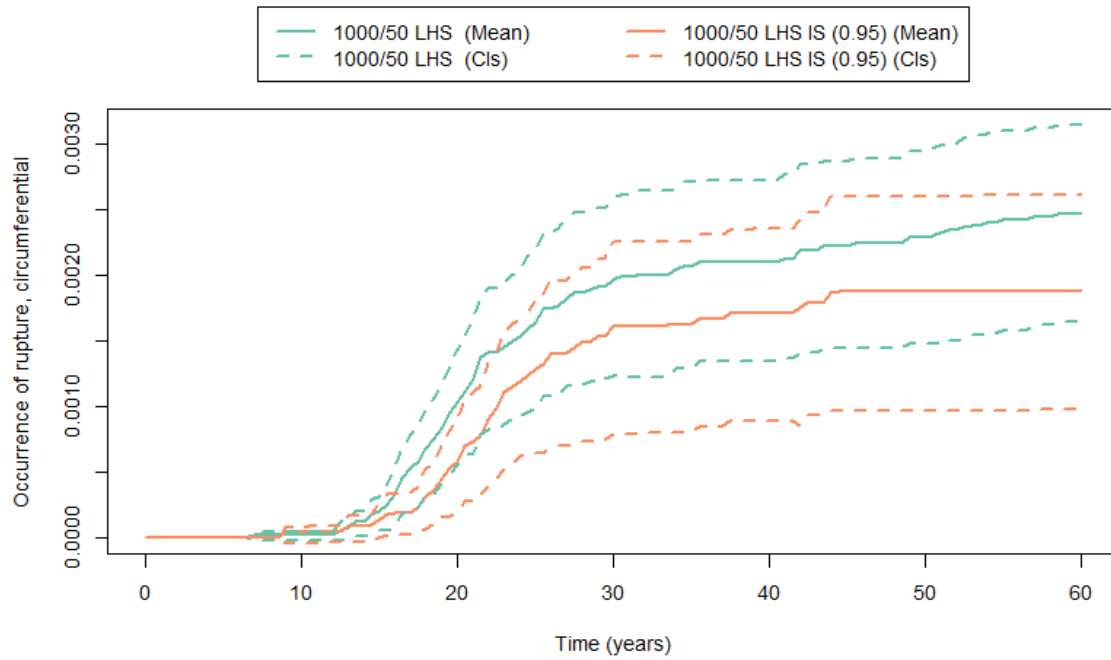


Figure 205: Mean probability of occurrence of circumferential ruptures (solid line) and 95% confidence intervals (dashed lines) for Scenario 10, Run 3 (green) and Run 4 (orange).

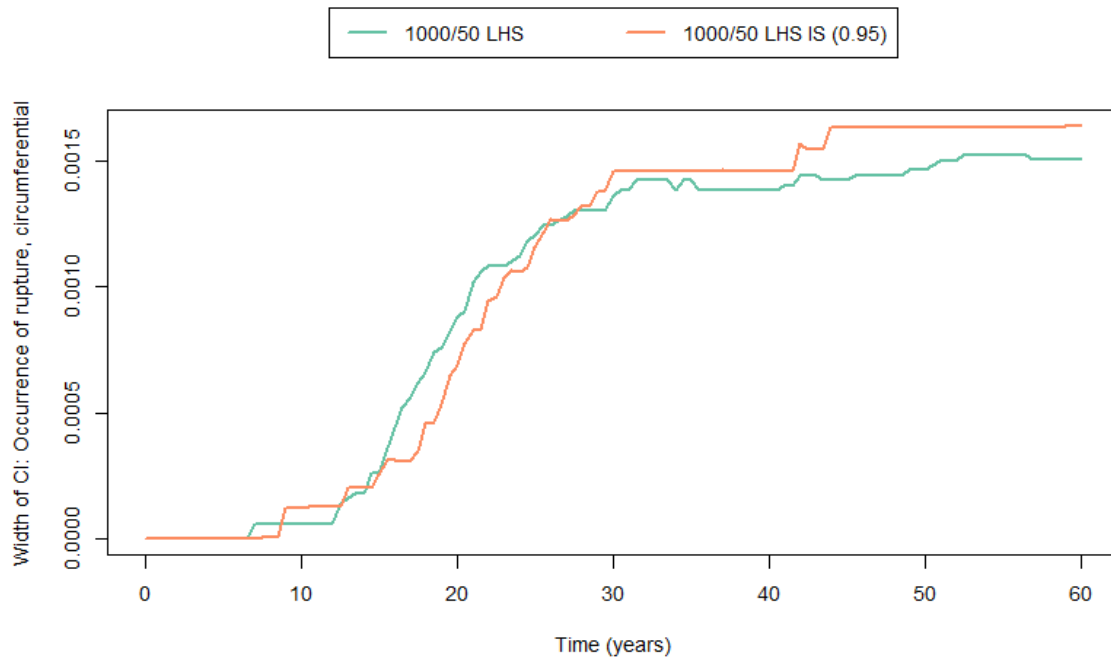


Figure 206: Width of 95% confidence interval for probability of occurrence of circumferential ruptures for Scenario 10, Run 3 (green) and Run 4 (orange).

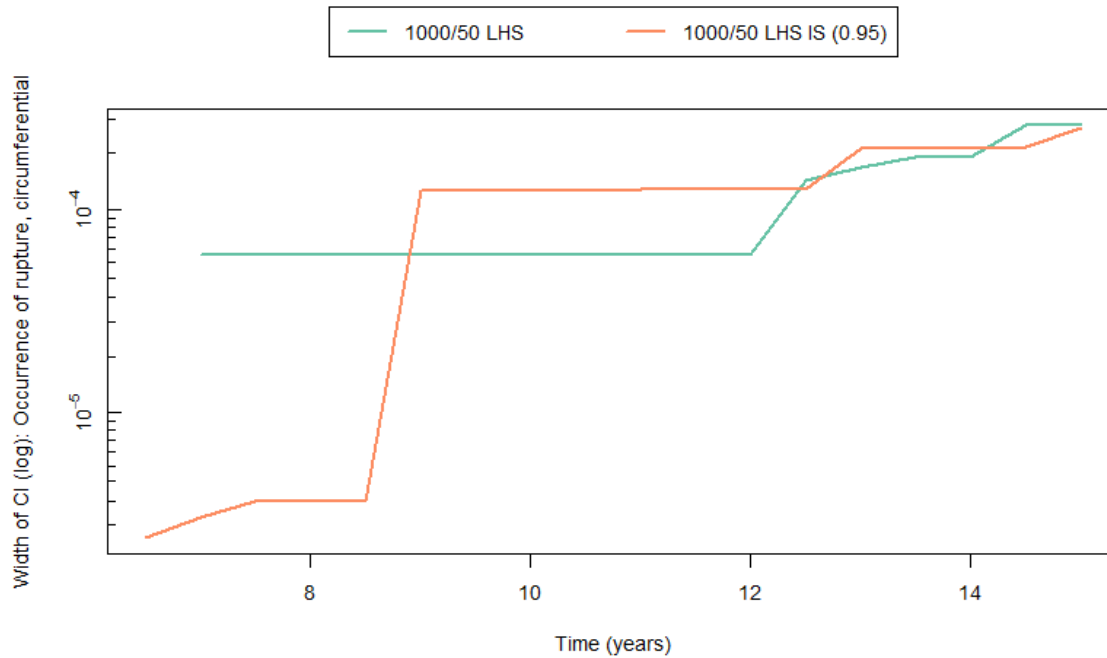


Figure 207: Width of 95% confidence interval for probability of occurrence of circumferential ruptures for Scenario 10, Run 3 (green) and Run 4 (orange) up to 15 years with y-axis shown in log scale.

3.10.3 Overall Results

The impact of sampling options selected for the analysis of Scenario 10 vary based on the output of interest. Figure 208 shows the mean probability of occurrence of circumferential rupture for all five of the sampling schemes used in the analysis of Scenario 10. While the options using importance sampling and/or LHS have mean values that are relatively close to one another, they are not nearly as well converged at the mean values of probability of occurrence of circumferential cracks shown in Figure 209. The probability of occurrence of circumferential ruptures is much lower than the probability of occurrence of circumferential cracks. This difference makes the probability of circumferential ruptures much more difficult to estimate. If a more accurate bound on this probability is required, additional epistemic realizations would help to alleviate this result spread.

The analysis of the epistemic uncertainty for the sampling scheme used in Run 4 (1000 epistemic, 50 aleatory realizations with LHS and IS on p2543) will be described in Section 3.10.4. This sampling scheme will also be examined using the convergence analysis techniques described in Section 2.4. The results of this convergence analysis are given in Section 3.10.5.

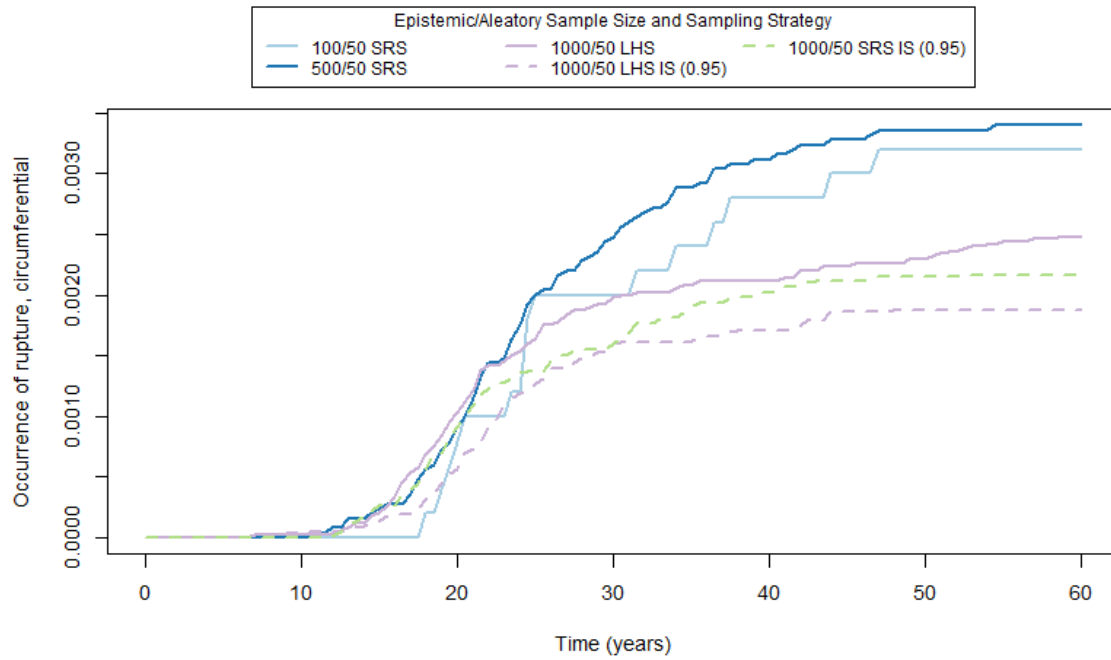


Figure 208: Mean probability of occurrence of circumferential crack ruptures for Scenario 10, Runs 1 through 5.

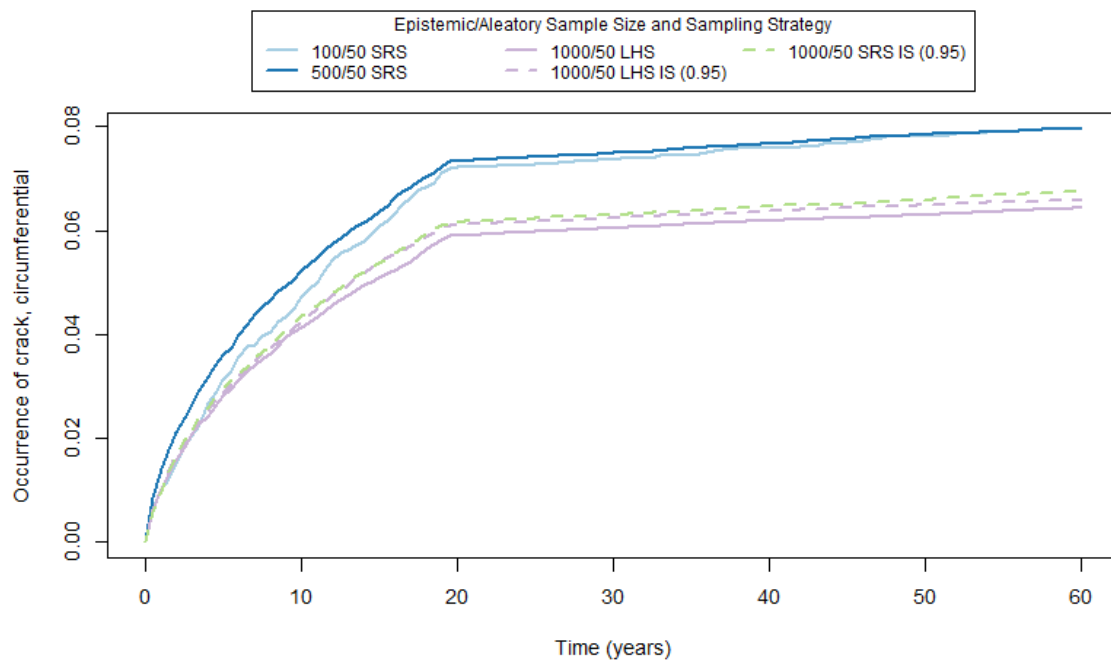


Figure 209: Mean probability of occurrence of circumferential cracks for Scenario 10, Runs 1 through 5.

3.10.4 Epistemic Uncertainty Analysis

An epistemic uncertainty analysis was conducted for the sampling scheme used in Run 4 (1000 epistemic, 50 aleatory realizations with LHS and IS on p2543).

Plots of the probabilities of occurrence of circumferential rupture and cracks are shown in Figure 210 and Figure 211, respectively. These figures show the probability of occurrence for each epistemic realization (grey lines), the mean (red line), and the 5th, 50th, and 95th percentiles (blue, green, and purple lines, respectively).

Because probability of occurrence of circumferential ruptures is small and, therefore, difficult to estimate, the difference in lower and upper quantiles is small and, at early time, the 5th, 50th, and 95th percentiles are indistinguishable from one another (Figure 210). The 95th percentile of probability of occurrence of circumferential ruptures for this run is approximately 2%, meaning that only 5% of the epistemic realizations have a greater than 2% chance of experiencing a circumferential rupture.

In contrast, the epistemic realizations show a greater spread in the estimation of the probability of occurrence of circumferential cracks, as shown in Figure 211. The 95th percentile of probability of occurrence of circumferential cracks for this run is approximately 35%, meaning that only 5% of the epistemic realizations have a greater than 35% chance of experiencing a circumferential crack.

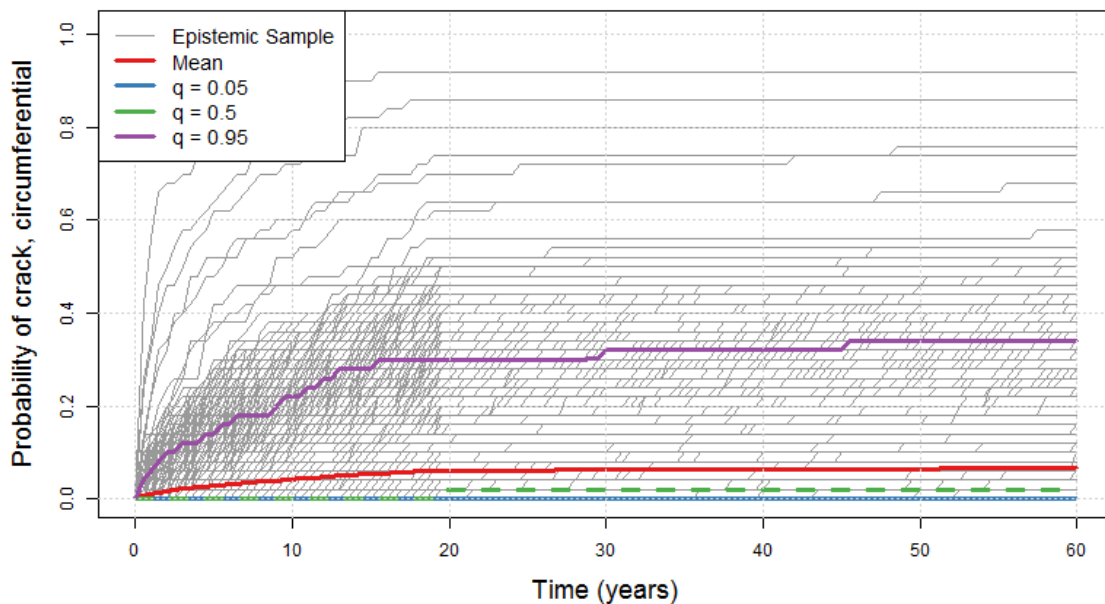


Figure 210: Probability of occurrence of circumferential rupture for each epistemic realization (grey), the mean (red), and the 5th (blue), 50th (green), and 95th (purple) percentiles for Scenario 10, Run 4.

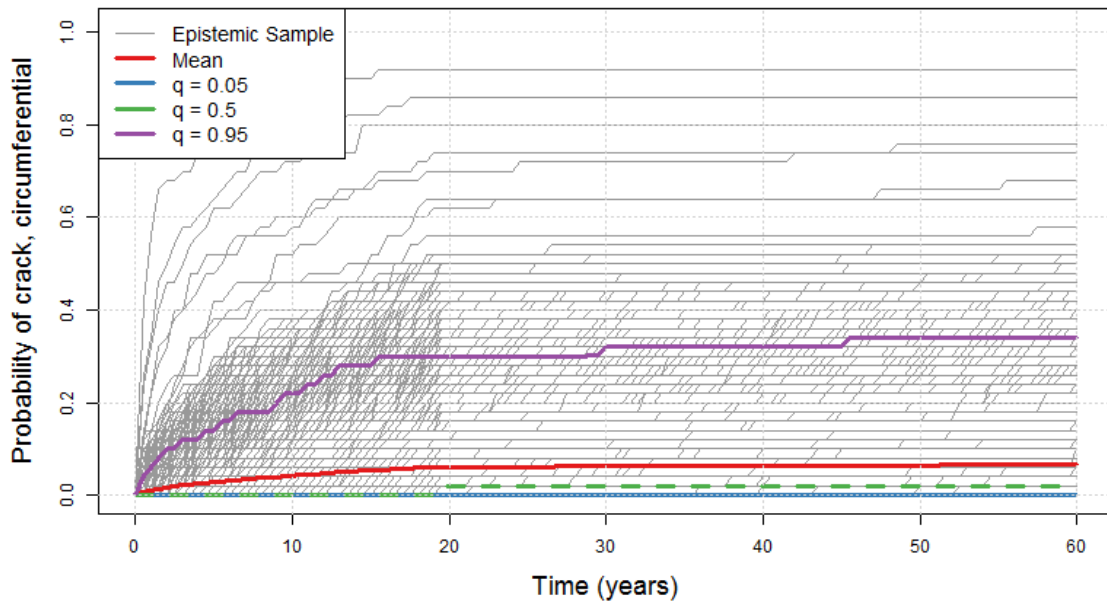


Figure 211: Probability of occurrence of circumferential cracks for each epistemic realization (grey), the mean (red), and the 5th (blue), 50th (green), and 95th (purple) percentiles for Scenario 10, Run 4.

The epistemic uncertainty results for probability of occurrence of circumferential leaks and ruptures are compared in Figure 212. As expected, circumferential leaks occur before circumferential ruptures and are more common. These leaks do not always lead to rupture, making the probability of circumferential ruptures more difficult to estimate as previously described.

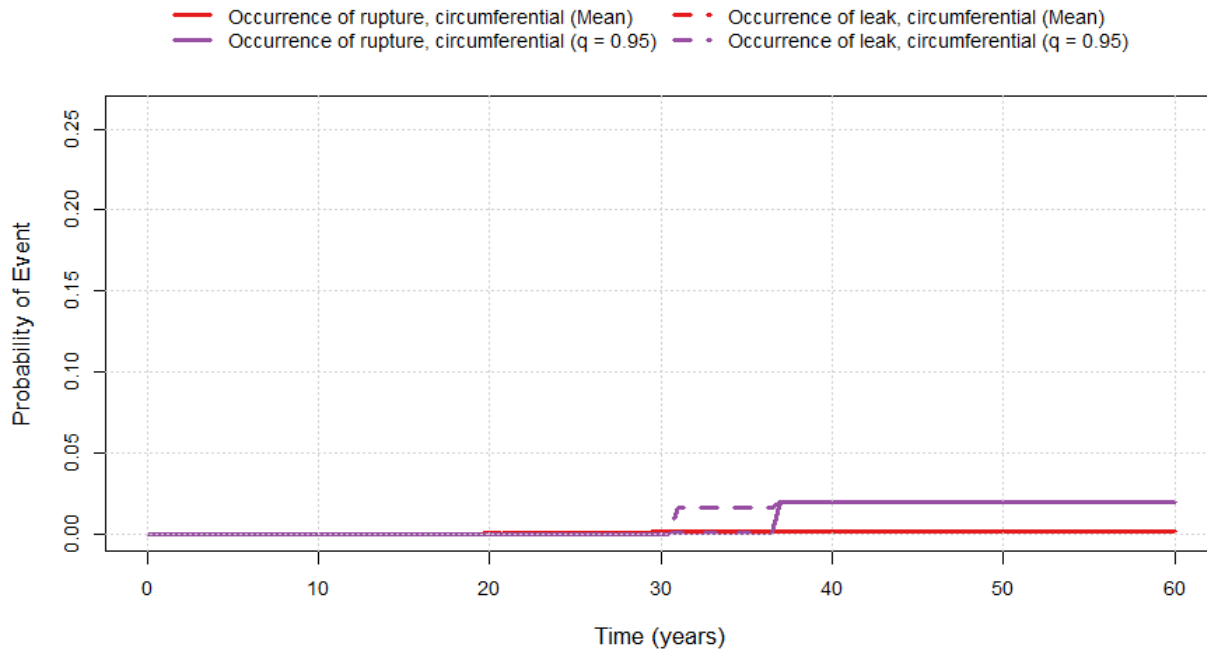


Figure 212: Mean (red) and 95th percentile (purple) for occurrence of circumferential rupture (solid) and crack (dashed) for Scenario 10, Run 4.

3.10.5 Convergence Analysis

The convergence of the sampling scheme used for Run 4 (1000 epistemic, 50 aleatory realizations with LHS and IS on p2543) for Scenario 10 was analyzed using five replicate simulations with different random seeds. Each replicate was used to give an estimate of the mean for each output of interest. The mean probability of occurrence of circumferential ruptures for each replicate is shown in Figure 213 and the mean probability of occurrence of circumferential cracks is shown in Figure 214. The replicates were also used to construct a 95% prediction interval about the mean probability of occurrence of circumferential ruptures and cracks, as shown in Figure 215 and Figure 216. The prediction intervals for probability of occurrence of circumferential ruptures have a width of approximately 0.3% while the prediction intervals for probability of occurrence of axial cracks have a width of 0.5%. The small magnitude of these intervals shows that the results are converging under the selected sampling scheme although, if smaller prediction intervals are required, additional epistemic realizations may be needed.

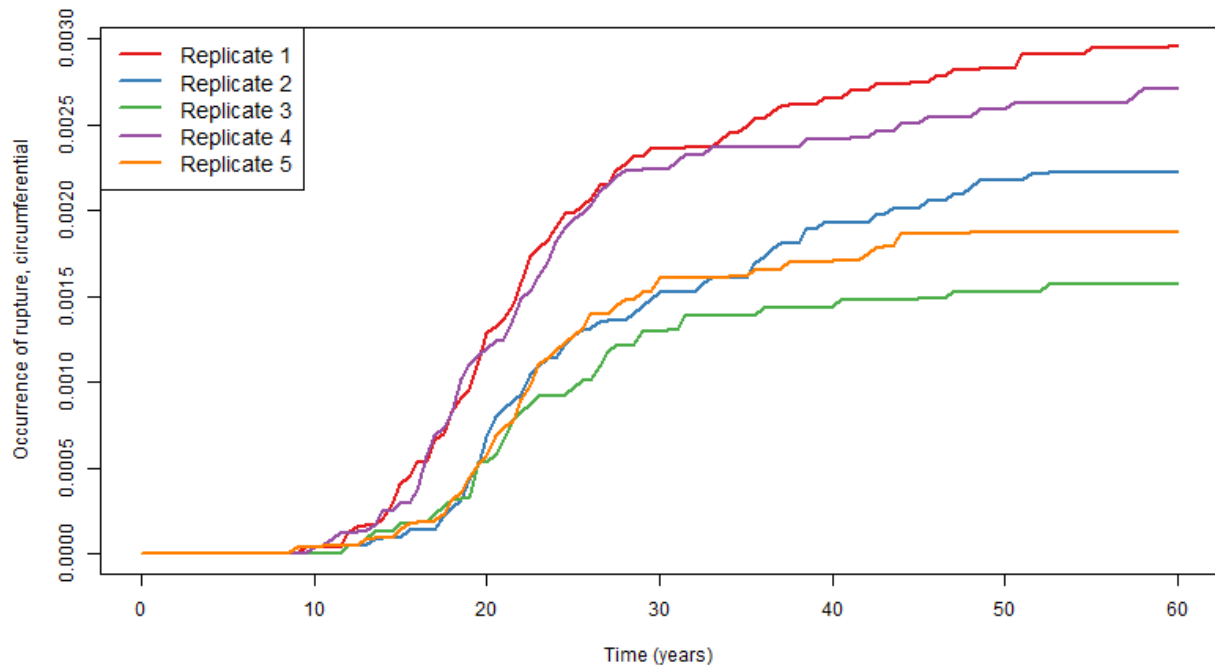


Figure 213: Comparison of estimated mean probability of occurrence of circumferential rupture for five convergence replicates for Scenario 10.

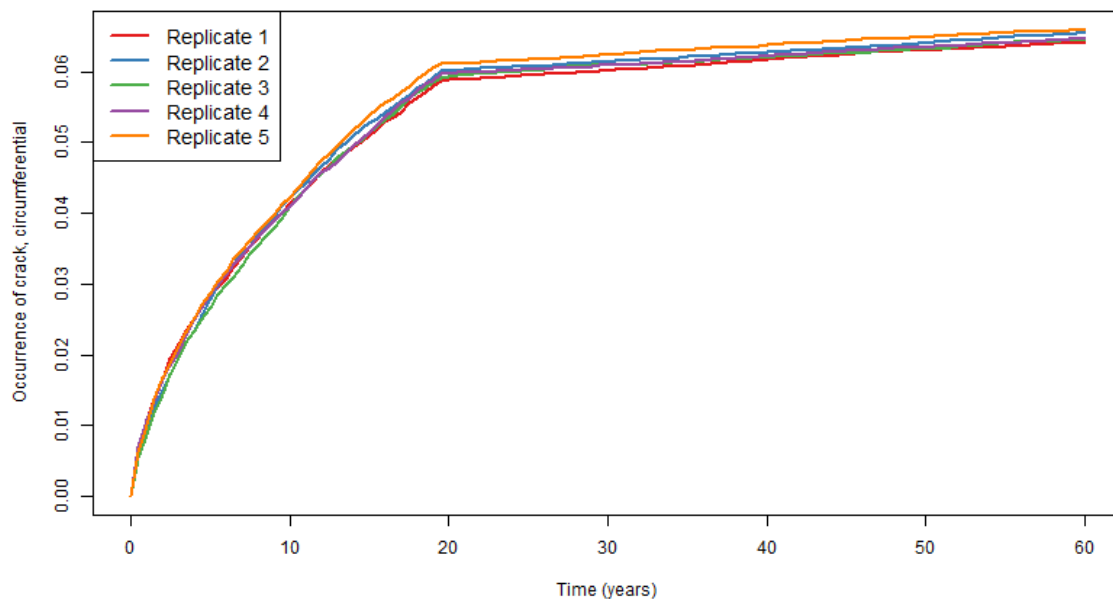


Figure 214: Comparison of estimated mean probability of occurrence of circumferential cracks for five convergence replicates for Scenario 10.

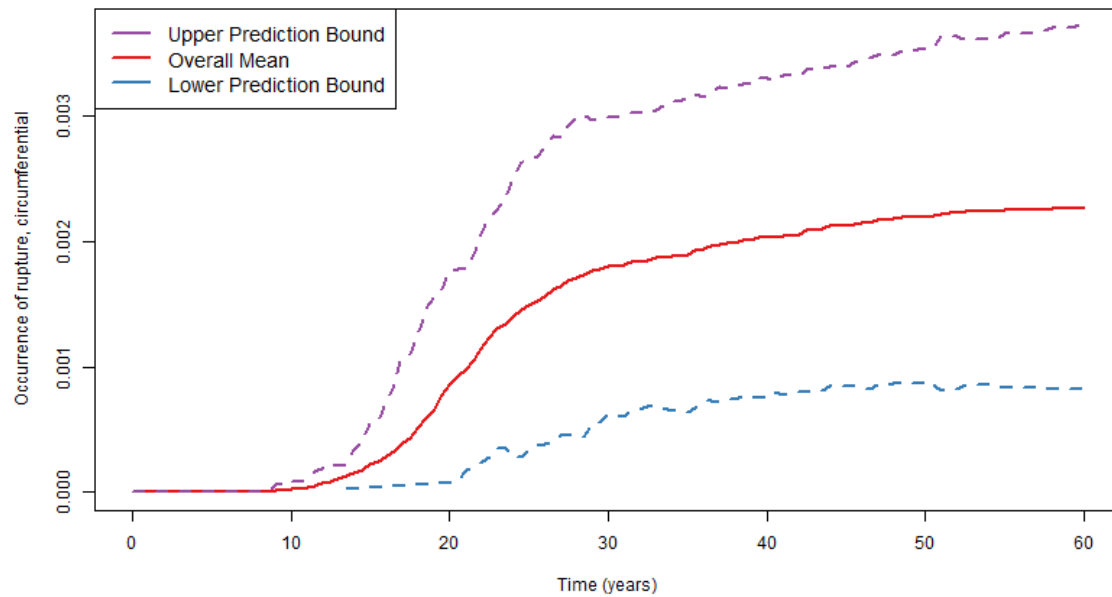


Figure 215: 95% prediction interval over mean probability of occurrence of circumferential ruptures for five replicates for Scenario 10.

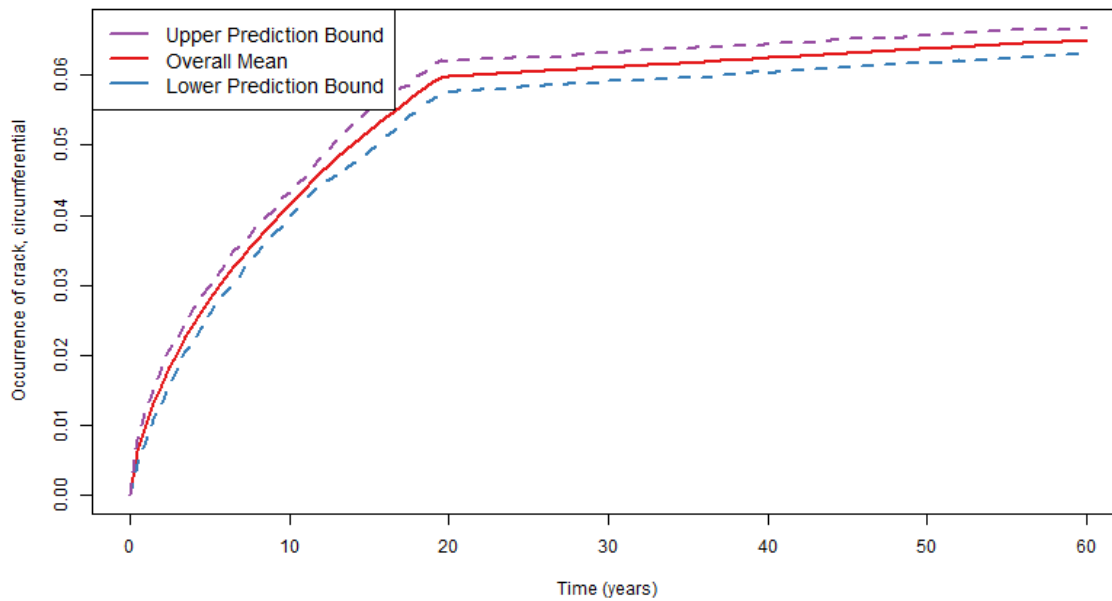


Figure 216: 95% prediction interval over mean probability of occurrence of circumferential cracks for five replicates for Scenario 10.

3.10.6 Comparison to Scenario 3, 4, and 8 Results

The results of Scenario 10, Run 4 were compared to Scenario 3, Run 9 results derived under the same sampling selections. These comparisons are shown for probability of occurrence of circumferential and axial cracks, leaks, and ruptures in Figure 217, Figure 218, and Figure 219. This comparison provides insight into the impact of the mitigation options prescribed for Scenario 10. These comparisons show that the mitigation options exercised in Scenario 10 cause a greater than 5% absolute decrease in the mean probability of occurrence of axial cracks and an approximately 3% decrease for circumferential cracks (Figure 217). The mitigation impact is also strong for probability of occurrence of leaks and ruptures. Scenario 10 sees an approximate 6.5% absolute decrease in the probability of occurrence of axial leaks and an approximate 2% absolute decrease in the probability of occurrence of circumferential leaks (Figure 218). The probability of occurrence of circumferential ruptures decreases by approximately 1.5% for Scenario 10 when compared to Scenario 3 (Figure 219). Axial ruptures were not seen in either scenario.

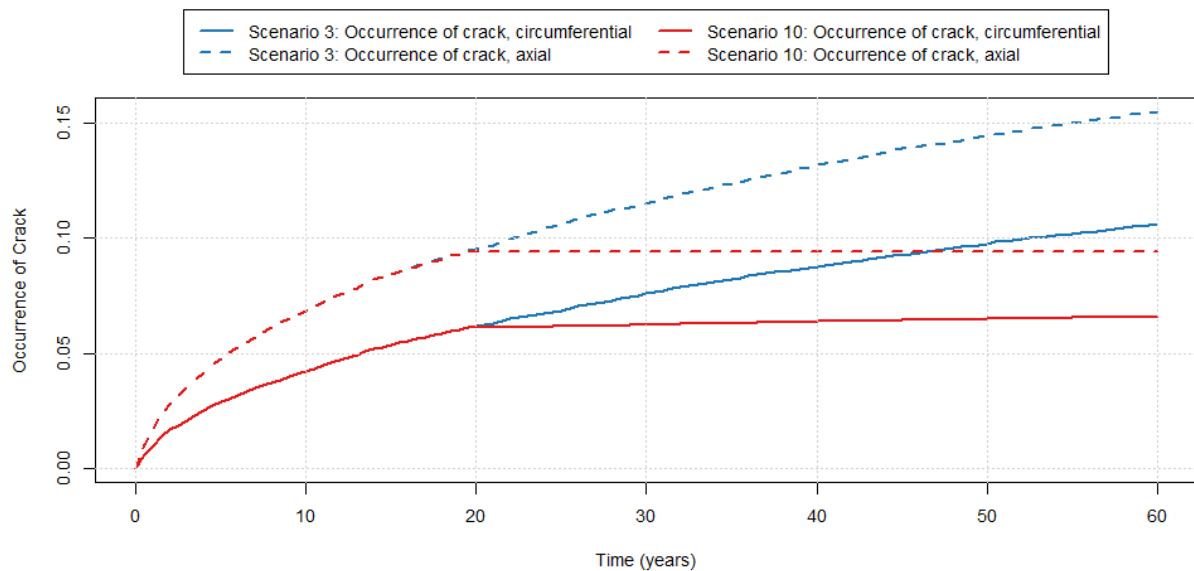


Figure 217: Mean probability of axial (dashed lines) and circumferential (solid lines) crack occurrence for Scenario 3 (blue lines) and Scenario 10 (red lines).

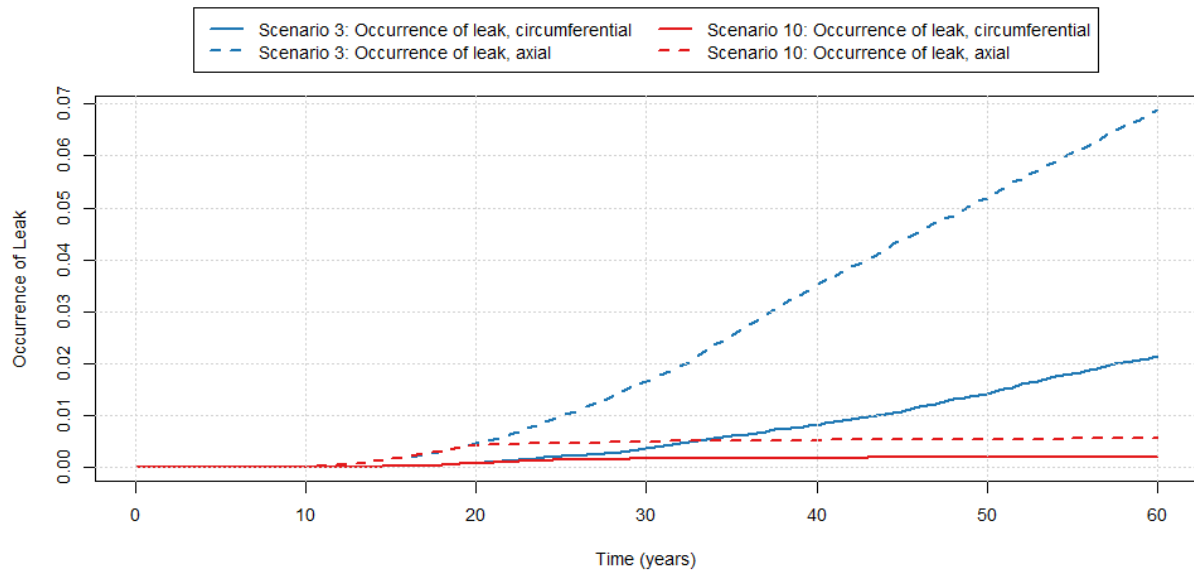


Figure 218: Mean probability of axial (dashed lines) and circumferential (solid lines) leak occurrence for Scenario 3 (blue lines) and Scenario 10 (red lines).

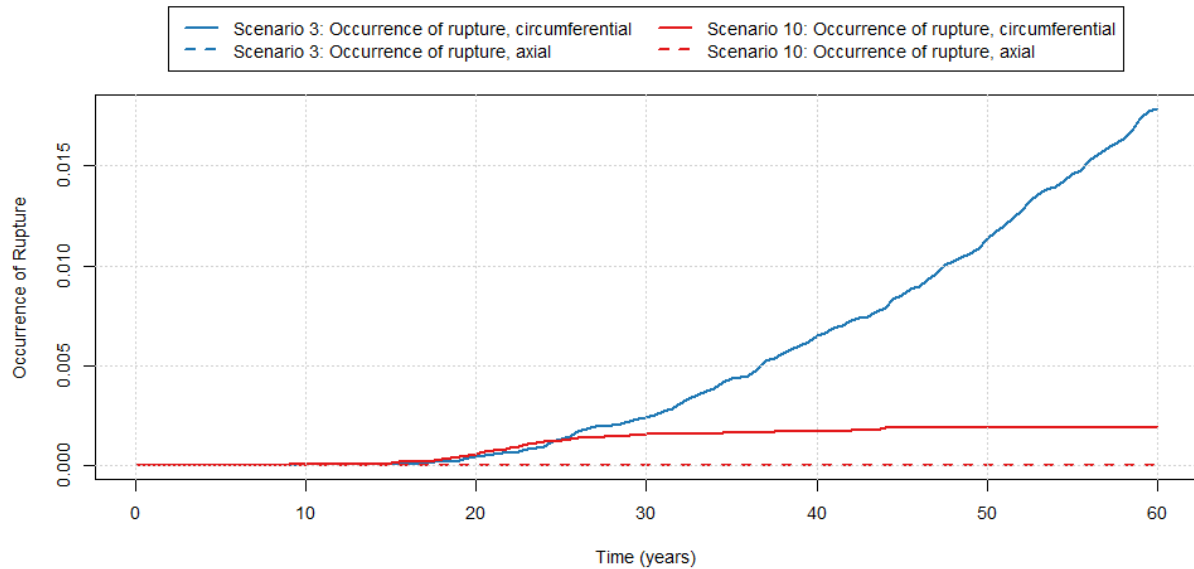


Figure 219: Mean probability of axial (dashed lines) and circumferential (solid lines) rupture occurrence for Scenario 3 (blue lines) and Scenario 10 (red lines).

These results can also be visually compared to those for Scenario 4 (Run 4) and Scenario 8 (Run 4) found in Sections 3.4.6 and 3.8.6, respectively. Scenario 4 also uses Scenario 3 as a basis but studies the impact of MSIP mitigation at 20 years. The occurrences of circumferential and axial cracks are compared for Scenarios 4 and 10 in Figure 220. Scenario 8 again uses Scenario 3 as a basis and studies the impact of both Zn and H₂ mitigation at 20 years. The occurrences of circumferential and axial cracks are compared for Scenarios 8 and 10 in Figure 221. Comparing these results shows that

the impact of MSIP mitigation dominates the Scenario 10 results as these results are very similar to those found in the analysis of Scenario 4. Slightly greater decreases can be found in the occurrence of cracks, leaks, and ruptures for Scenario 10, indicating that the chemical mitigation options analyzed in Scenario 8 do have some impact.

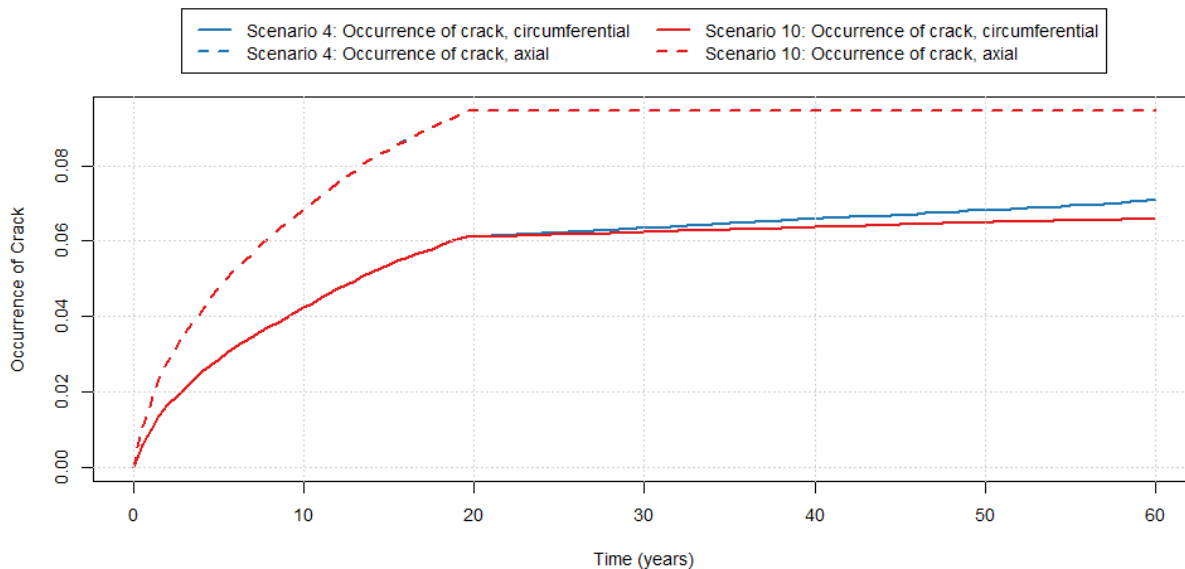


Figure 220: Mean probability of axial (dashed lines) and circumferential (solid lines) crack occurrence for Scenario 4 (blue lines) and Scenario 10 (red lines).

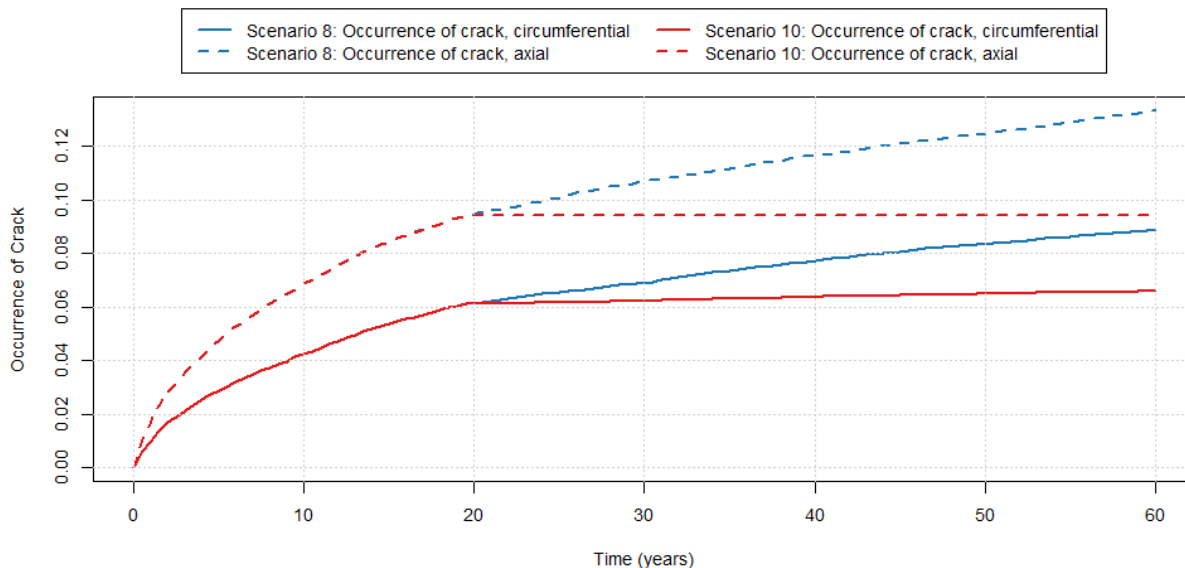


Figure 221: Mean probability of axial (dashed lines) and circumferential (solid lines) crack occurrence for Scenario 8 (blue lines) and Scenario 10 (red lines).

The inclusion of fatigue initiation and growth in Scenario 10 does not have a strong impact on the results when compared to the PWSCC-only initiation and growth used in Scenario 3. The Scenario

10 results are nearly identical to those of Scenario 3 for all outputs up to the 20-year mitigation time indicating that the same crack initiation and growth mechanisms are dominating the results for these two scenarios.

3.11 Scenario 11

3.11.1 Scenario Summary

Scenario 11 examines fatigue flaw initiation on the same RPV outlet weld of the other scenarios. PWSCC initiation and growth are turned off so that cracks only initiate and grow by fatigue. The following options are defined:

- Fatigue flaw initiation
- Circumferential and axial flaw orientation
- Fatigue flaw growth
- No mitigation

The transients used for this scenario are the same as those used in Scenarios 1 and 10 (the other scenarios involving fatigue) except that they include an additional thermal transient. The additional transient does not have a technical basis and is purely hypothetical. It is used to initiate fatigue flaws; otherwise, no fatigue flaws occur and no sensitivity studies can be performed for this scenario. The additional transient initially cools the weld by 100°C. Then, over 5 seconds, it heats the weld to 100°C above the nominal operating temperature and holds it there for another 5 seconds. This transient is set to occur 305 times each year for the entire plant operation period. Among the epistemically varying input parameters, the occurrence of flaws is sensitive to the fatigue initiation load sequence factor, p2522.

Leaks and ruptures do not occur in this scenario. Therefore, the outputs of interest defined for this scenario are simply:

- Probability of occurrence of axial crack
- Probability of occurrence of circumferential crack

The simulations performed for this scenario are listed below.

1. 100/50 (epistemic/aleatory sample size) with SRS
2. 500/50 with SRS
3. 1000/50 with LHS
4. 1000/50 with LHS and importance sampling on p2522 (fatigue initiation load sequence factor, FLOAD) with target quantile of 0.95
5. 1000/50 with SRS and IS on p2522 with target quantile of 0.95
6. 1000/50 with LHS and importance sampling on p2522 (fatigue initiation load sequence factor, FLOAD) with target quantile of 0.995
7. 1000/50 with SRS and IS on p2522 with target quantile of 0.995

In the Scenario 11 analysis, only results of circumferential cracks are presented and discussed. The circumferential crack results are chosen because the results of circumferential cracks are presented for Scenario 3 (Section 3.3). The occurrences of axial and circumferential cracks are similar for Scenario 11.

Of the various sampling schemes tested, the lowest uncertainty in the mean probability of circumferential crack occurrence at early times is achieved with 1000 epistemic realizations sampled using LHS and importance sampling (Section 3.11.2). The fatigue initiation load sequence

factor (p2522) is the only significant parameter of the epistemically-varying parameters in this scenario that affects fatigue crack initiation. The number of aleatory samples is fixed at 50 for all simulations.

Scenario 11 results in an overall probability of circumferential crack occurrence of 0.001 by year 21, 0.01 by year 39, and 0.04 by year 60 (Section 3.11.4). Convergence testing indicates that results are stable and reproducible regardless of random seeds (Section 3.11.5).

3.11.2 Analysis Progression

Runs 1 and 2 assess the effects of epistemic sample size on result uncertainty. Figure 222 plots the mean probability of occurrence of circumferential cracks for Run 1 and Run 2, along with the 95% CIs, which describe the uncertainty associated with the mean. The means are similar for these runs, but the CI is narrower for the 100/50 SRS run (Run 1). The widths of the CIs are directly compared in Figure 223. At 60 years, the width of the CI for Run 1 is approximately 0.8 of the width of the CI for Run 2. This suggests that an increase from 100 to 500 epistemic realizations causes increased uncertainty in the mean number of circumferential crack occurrences for this scenario. This result is unexpected and is presumed due to chance. It is thought that Run 1 by chance did not realize enough of the uncertainty in the outputs due to too few epistemic realizations. The 99% confidence intervals, however, are reverse as shown in Figure 224 and Figure 225, as would be expected.

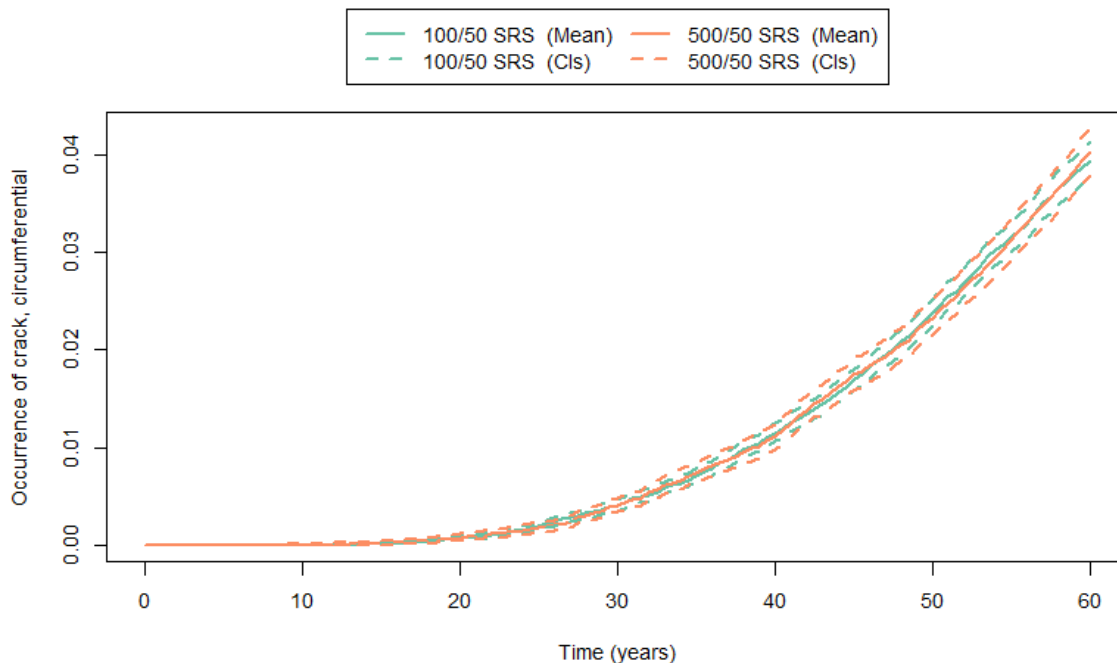


Figure 222: Mean probability of occurrence of circumferential cracks (solid line) and 95% confidence intervals (dashed lines) for Scenario 11, Run 1 (green) and Run 2 (orange).

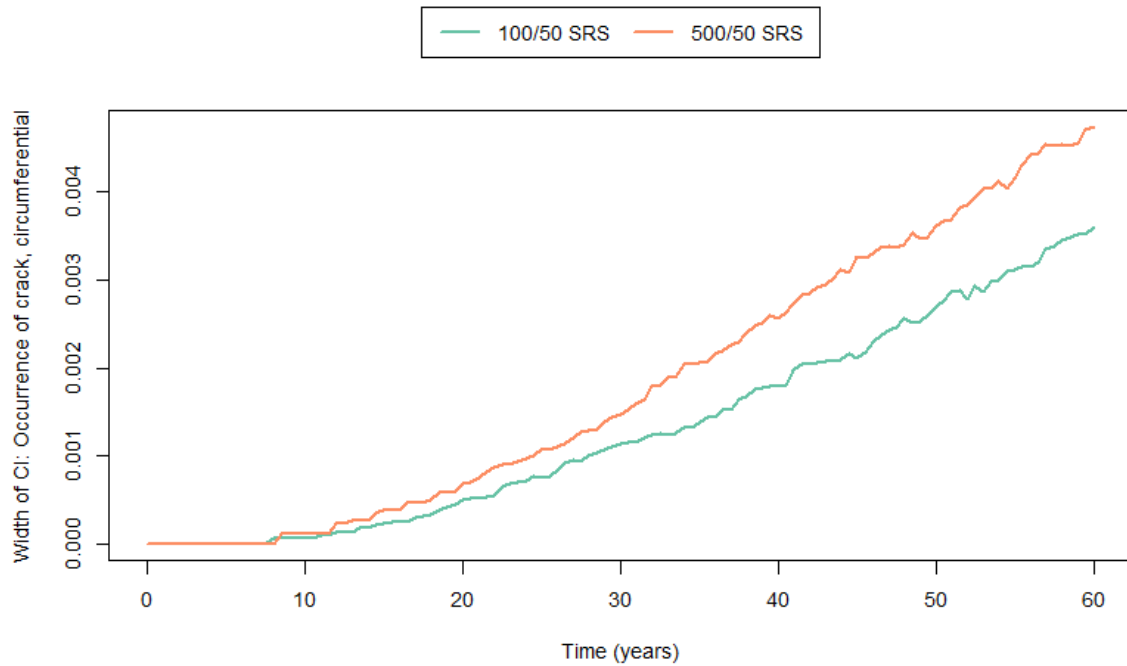


Figure 223: Width of 95% confidence interval for probability of occurrence of circumferential cracks for Scenario 11, Run 1 (green) and Run 2 (orange).

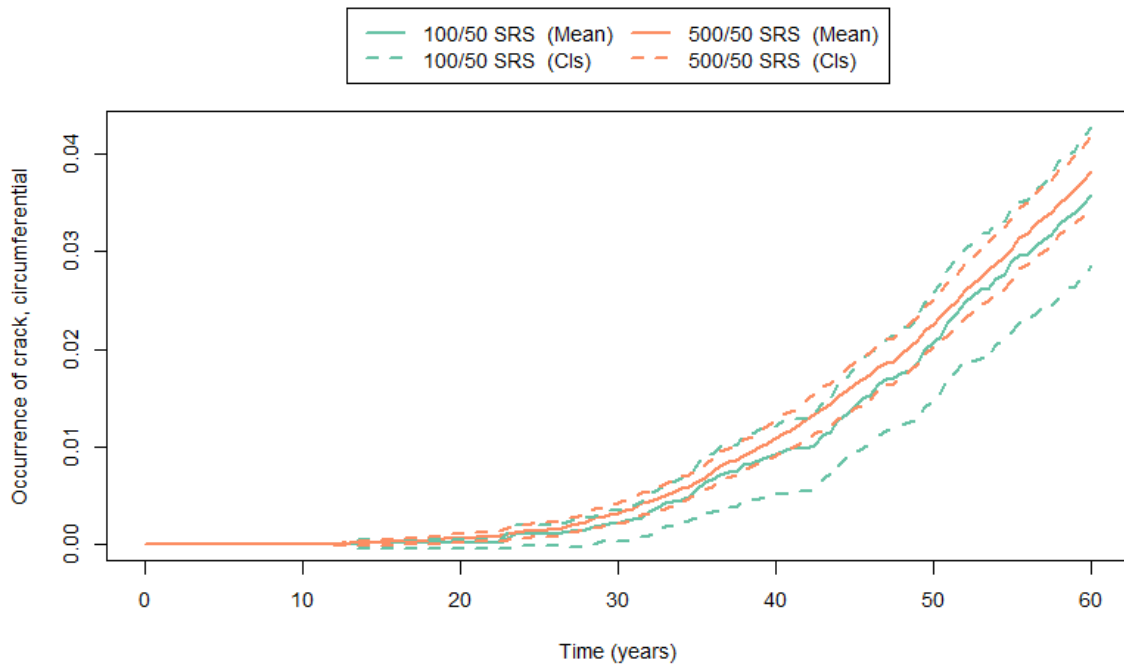


Figure 224: Mean probability of occurrence of circumferential cracks (solid line) and 99% confidence intervals (dashed lines) for Scenario 11, Run 1 (green) and Run 2 (orange).

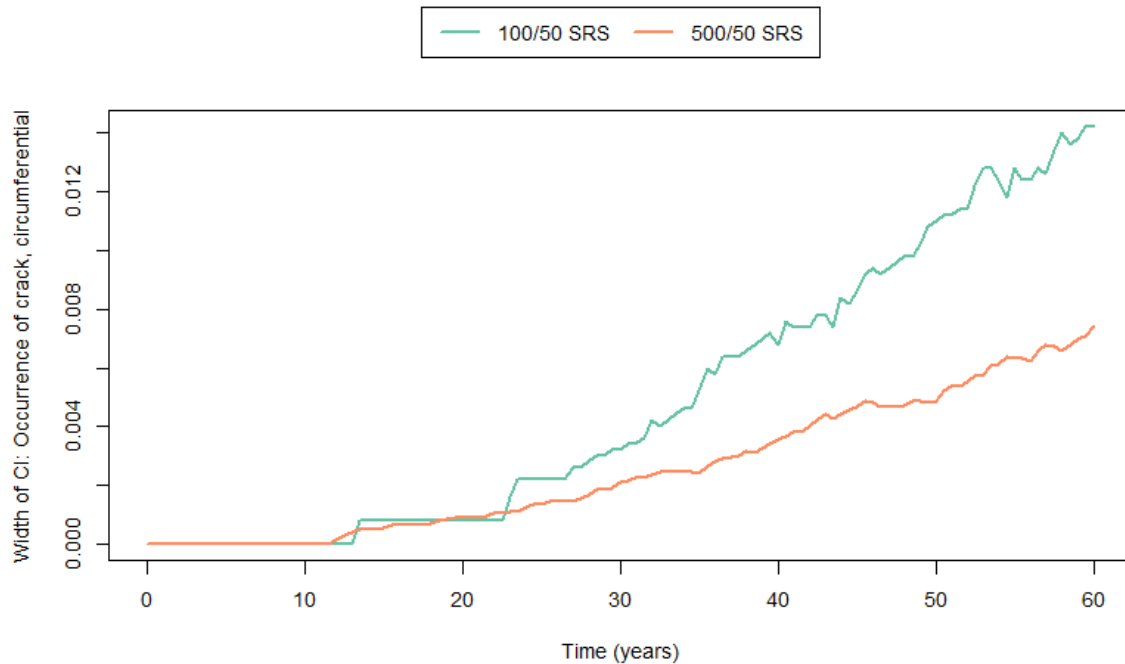


Figure 225: Width of 99% confidence interval for probability of occurrence of circumferential cracks for Scenario 11, Run 1 (green) and Run 2 (orange).

The effects of importance sampling are examined using Runs 3, 4, and 6. Each of these runs uses LHS. Run 3 has no importance sampling, while Runs 4 and 6 perform importance sampling on p2522 with target quantiles of 0.95 and 0.995, respectively.

Figure 226 shows the mean probability of circumferential crack occurrence as a function of time for Runs 3 and 6 along with 95% confidence intervals. A figure comparing the results of Runs 3 and 4 would be almost identical. Importance sampling as implemented in xLPR V2 is designed to provide tighter confidence intervals for extremely low probability occurrences. Low probability of circumferential crack occurrence is observed in this scenario in the first 15 years. Figure 227 shows that the confidence intervals are much tighter in the first 12 years using the importance sampling of Run 6 as compared to no importance sampling. However, at later times where probabilities are not a low, confidence intervals are wider with importance sampling. Figure 228 indicates that the target quantile, whether 0.95 or 0.995, makes little difference in the size of the confidence interval in this example.

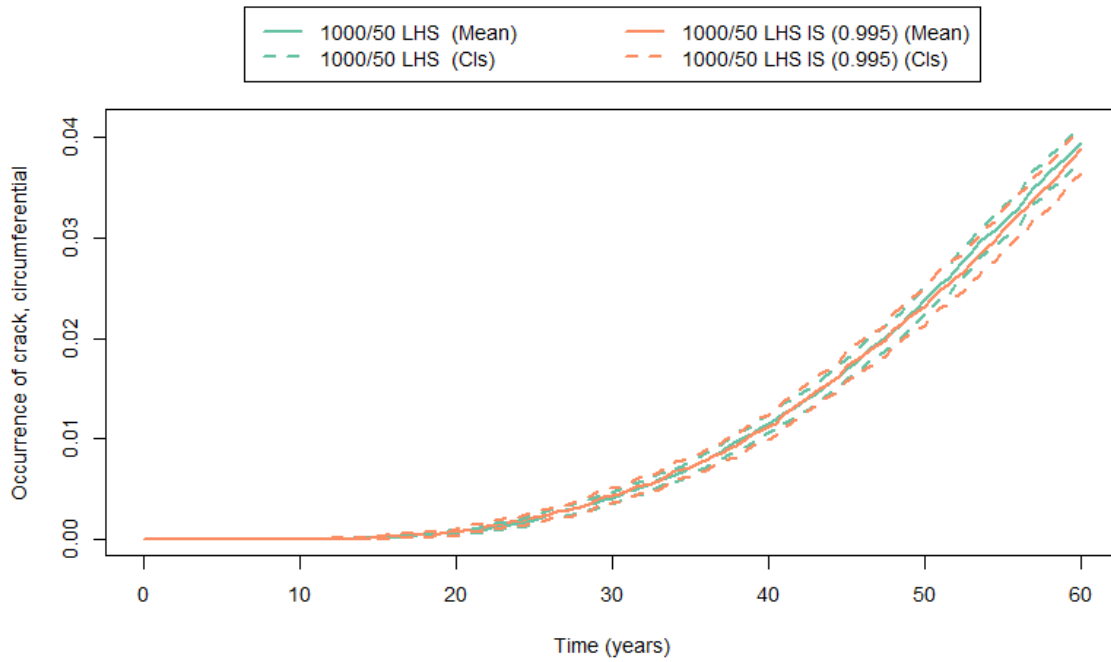


Figure 226: Mean probability of occurrence of circumferential cracks (solid line) and 95% confidence intervals (dashed lines) for Scenario 11, Run 3 (green) and Run 6 (orange).

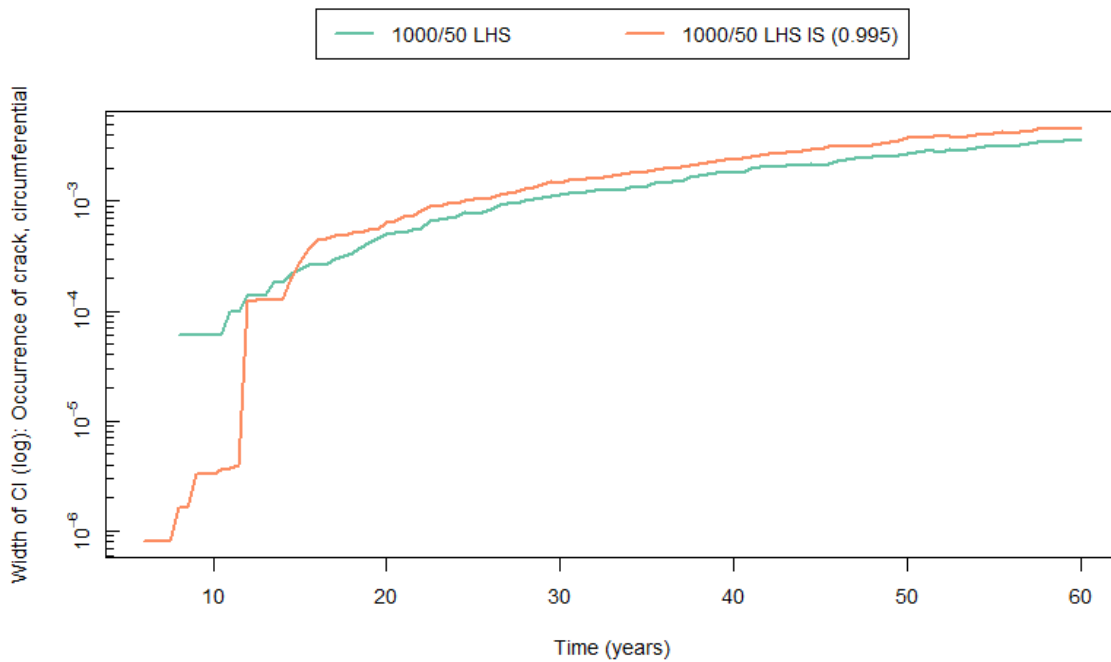


Figure 227: Width of 95% confidence interval for probability of occurrence of circumferential cracks for Scenario 11, Run 3 (green) and Run 6 (orange).

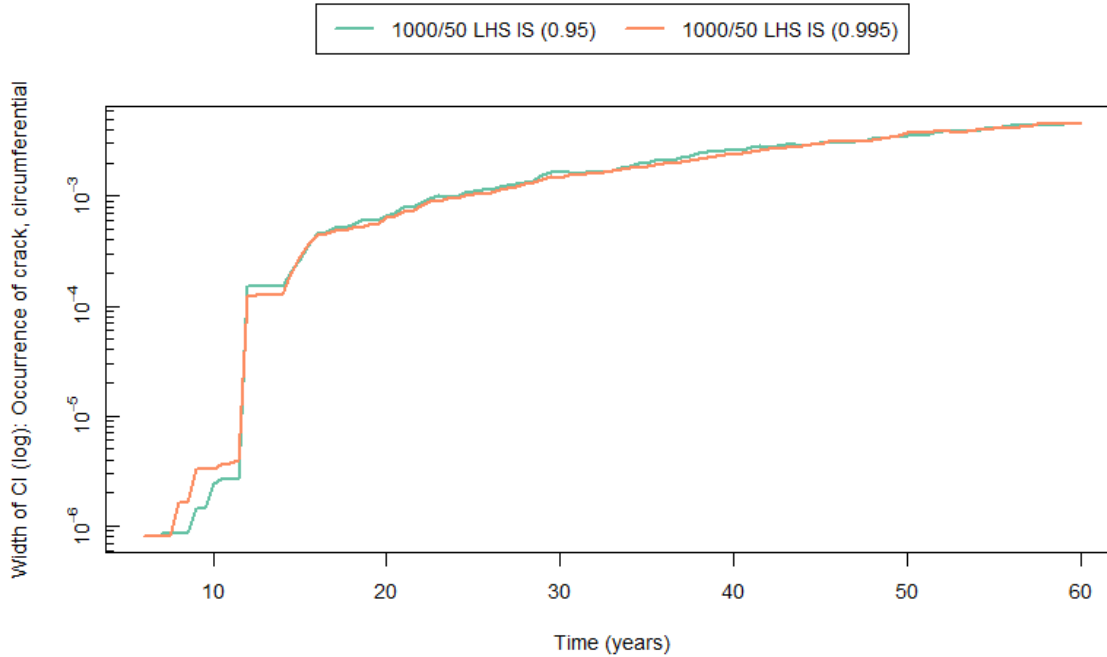


Figure 228: Width of 95% confidence interval for probability of occurrence of circumferential cracks for Scenario 11, Run 4 (green) and Run 6 (orange).

The effects of importance quantiles 0.95 and 0.995 when using SRS are shown in Figure 231 and Figure 232. The mean probability and confidence intervals are practically identical in Figure 231. The widths of the confidence intervals are also nearly the same except on log scale at early times. Compared to the LHS confidence interval widths at early times (Figure 228), the SRS widths are not calculated until several years later.

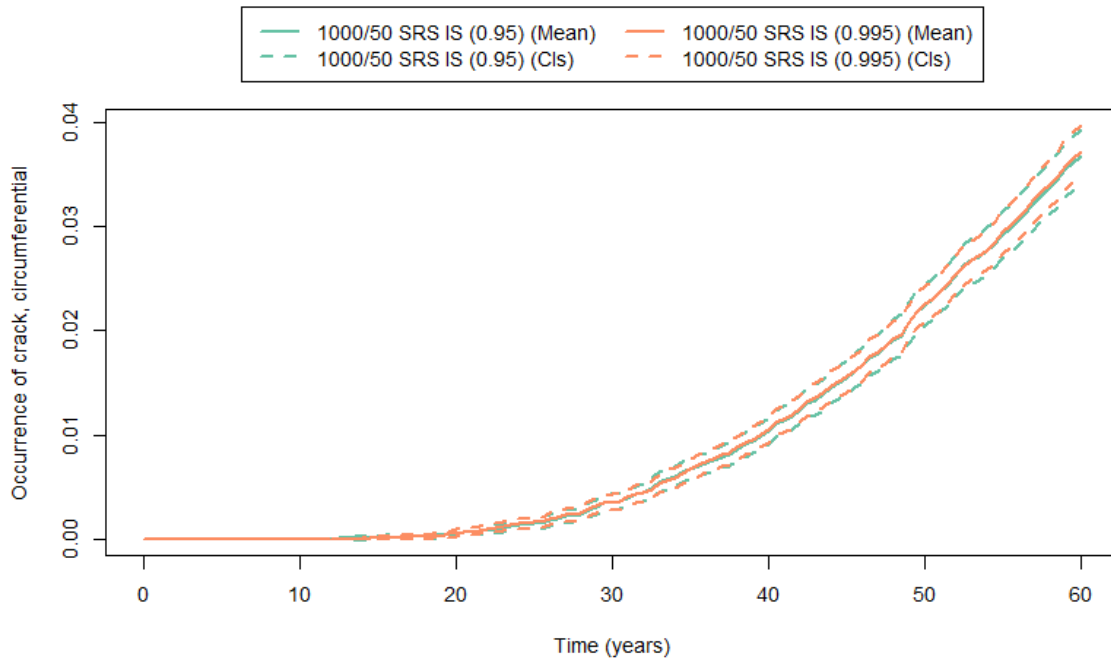


Figure 229: Mean probability of occurrence of circumferential cracks (solid line) and 95% confidence intervals (dashed lines) for Scenario 11, Run 5 (green) and Run 7 (orange).

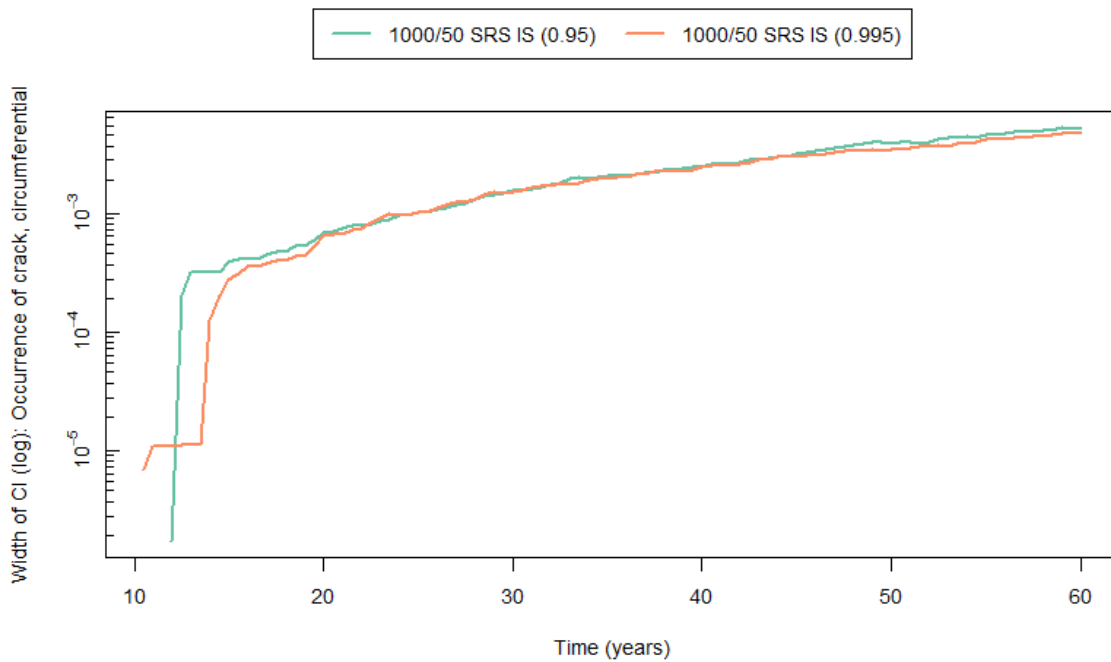


Figure 230: Width of 95% confidence interval for probability of occurrence of circumferential cracks for Scenario 11, Run 5 (green) and Run 7 (orange).

Of all epistemically-varying parameters that affect fatigue crack initiation, p2522 is the most sensitive parameter. Table 12 shows the top five parameters most correlated with the occurrence of circumferential cracks in this analysis. Only p2522 is significantly correlated. Based on model equations, none of the other variables listed in the table have any effect on crack occurrence; they only show slight correlation by chance. Scatterplots of the sampled p2522 values are shown in Figure 231. The scatterplot on the left is without importance sampling (Run 3). The scatterplots in the middle and on the right are with importance sampling at target quantiles of 0.95 (Run 4) and 0.995 (Run 6), respectively. The result stratification shown in Figure 231 is a result of the selected sample size. These results represent an average over a fixed number of indicator functions (value of 0 or 1). Thus, the probability can only occur incrementally as a function of the aleatory sample size.

Several parameters that are not defined with epistemic uncertainty (and therefore are not evaluated in the analysis presented in Table 12) may also be useful in importance sampling for fatigue initiation. One is likely the multiplier for the spatially-variable fatigue initiation strain threshold (STH), p2526. This parameter would be importance sampled on the low end, e.g., a target quantile of 0.005. Also, the multiplier for the spatially-variable fatigue initiation parameter C0, p2529, may be a good parameter for importance sampling (on the high end).

Table 12: Summary of the stepwise rank regression analysis for Scenario 11, Run 3 for circumferential crack occurrence by year 60.

Variable Identifier	Variable Name	SRRC
Probability of Occurrence of Circumferential Cracks – $R^2 = 0.354$		
p2522	Load Sequence Factor, FLOAD	0.572
p5103	Intercept, B0 (axial)	-0.095
p5106	b (circ)	-0.088
p5101	Beta_0 (circ)	-0.088
p9001	Fatigue Growth CKTH	-0.082

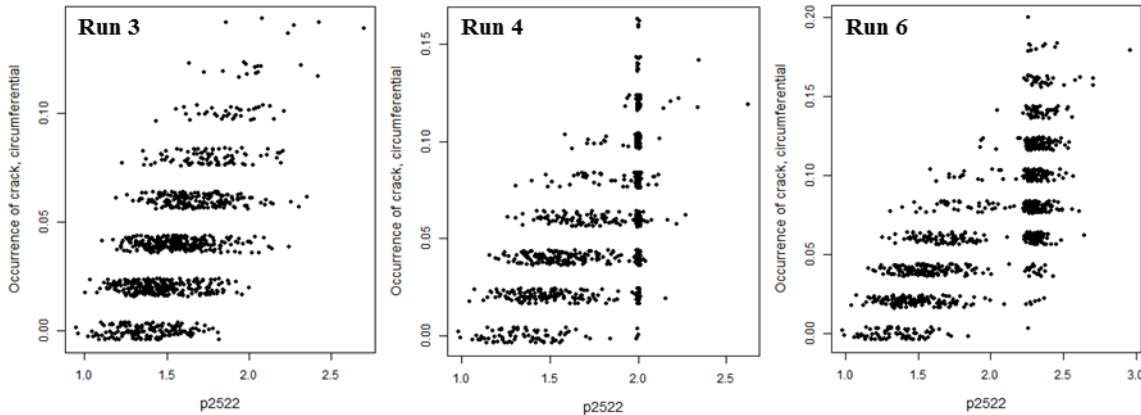


Figure 231: Scatterplots for variable p2522 (fatigue initiation load sequence factor, FLOAD) and the probability of occurrence for circumferential cracks for Scenario 11, Run 3, 4, and 6.

Run 7 is the same as Run 6 except that SRS is used instead of LHS. Both have importance sampling on p2522 with a target quantile of 0.995. As shown in Figure 232, the mean circumferential crack occurrence for Run 7 (SRS) is not as smooth and is generally slightly lower than that of Run 6 (LHS).

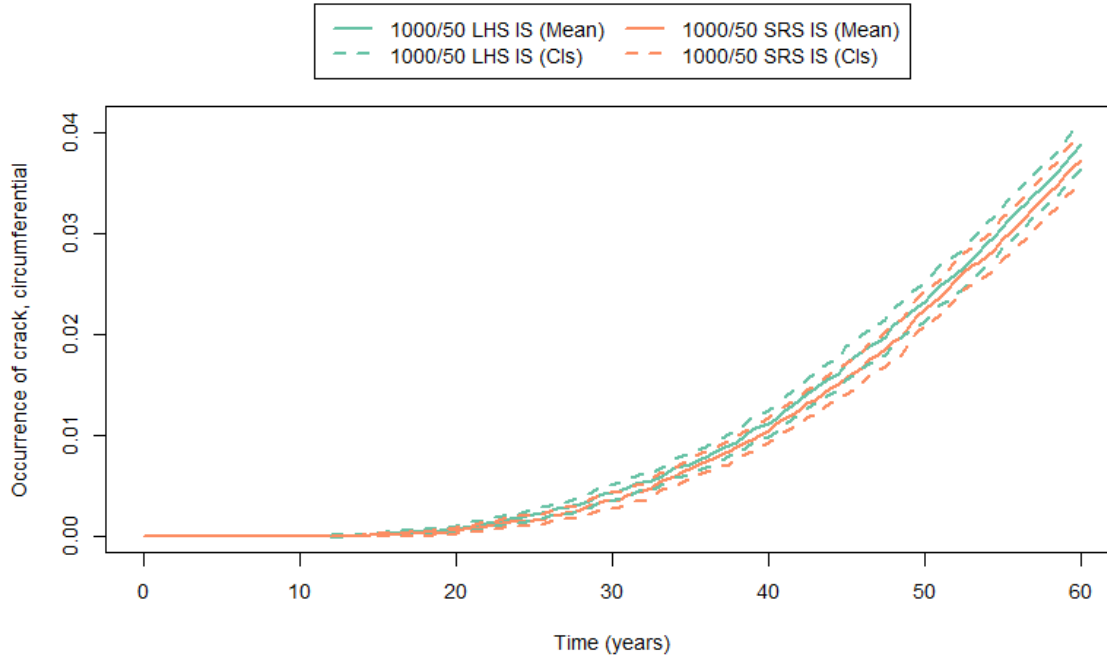


Figure 232: Mean probability of occurrence of circumferential cracks (solid line) and 95% confidence intervals (dashed lines) for Scenario 11, Run 6 (green) and Run 7 (orange).

3.11.3 Overall Results

The mean number of circumferential crack occurrences at 60 years generally converges as the number of epistemic realizations increases. This is indicated in Figure 233. The most significant outlier in Figure 233 is the 100/50 SRS curve, which is expected because it has the fewest number of realizations. The means of the other runs do not differ by much over the simulation period.

As for the effect of the number of epistemic realizations on confidence intervals of the mean number of circumferential crack occurrences for this scenario, the results are inconclusive. The confidence intervals significantly tighten at the 99% level as the number of epistemic realizations increases from 100 to 500 (Figure 225) but they widen at the 95% level (Figure 223). The remaining Scenario 11 simulations use LHS and/or importance sampling and therefore cannot be used to provide conclusive evidence of narrower confidence intervals with increasing numbers of epistemic simulations.

Significant reductions in confidence interval width are achieved with importance sampling at early times (Figure 227) and the use of LHS over SRS. However, at later times, when probabilities are much higher, importance sampling results in wider confidence intervals (Figure 227 and Figure 228).

The parameter chosen for importance sampling, p2522 (fatigue initiation load sequence factor), is significantly correlated with circumferential crack occurrence (Table 12 and Figure 231). The effects of importance sampling might be more pronounced if the number of epistemic realizations were limited to 100 or if the probability of circumferential crack occurrences were markedly lower.

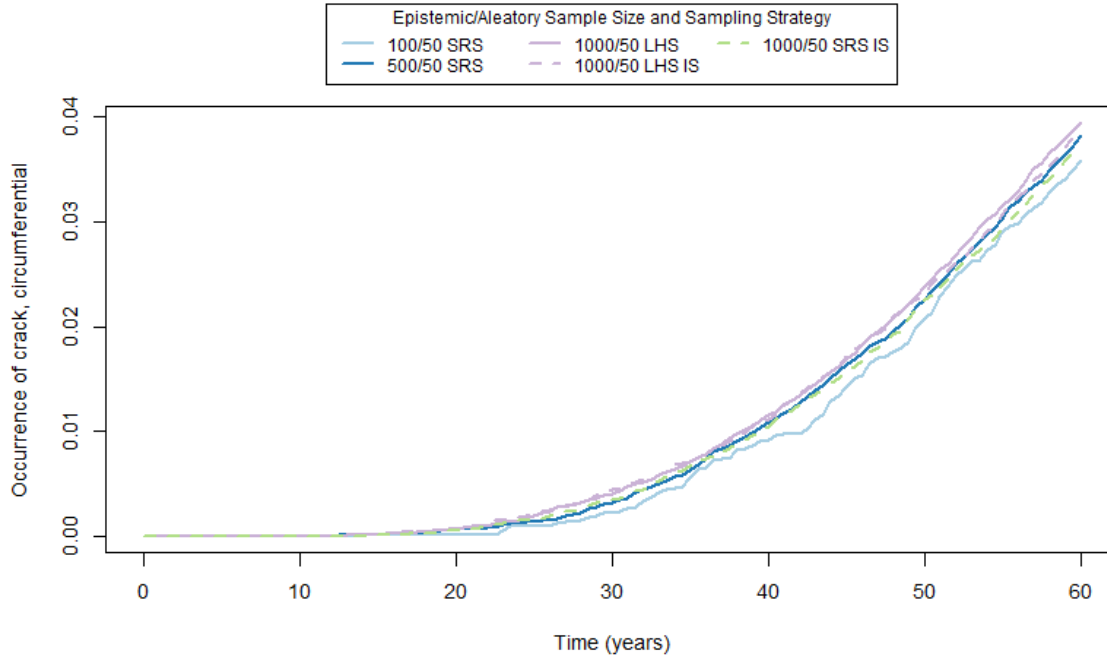


Figure 233: Mean probability of occurrence of circumferential cracks for Scenario 11, Runs 1, 2, 3, 6, and 7.

3.11.4 Epistemic Uncertainty Analysis

An examination of the spread of epistemic realizations and extreme quantiles is shown in Figure 234 for Scenario 11 Run 6 (1000/50 LHS IS at 0.995). Figure 234 shows a plot of the probability of circumferential crack occurrence for each epistemic realization (grey), the weighted mean (red), and the 5th (blue), 50th (green), and 95th (purple) percentiles over time. The mean and percentile curves account for the variable epistemic weights resulting from importance sampling. The realization curves indicate that the maximum observed frequency of circumferential crack occurrence by year 60 is 0.20. The 95th percentile curve indicates a 5% probability by year 22 of an epistemic realization with at least one circumferential crack occurrence; the 50th percentile curve indicates a 50% probability by year 43. The mean data points indicate an overall probability of circumferential crack occurrence of 0.001 by year 21, 0.01 by year 39, and 0.04 by year 60.

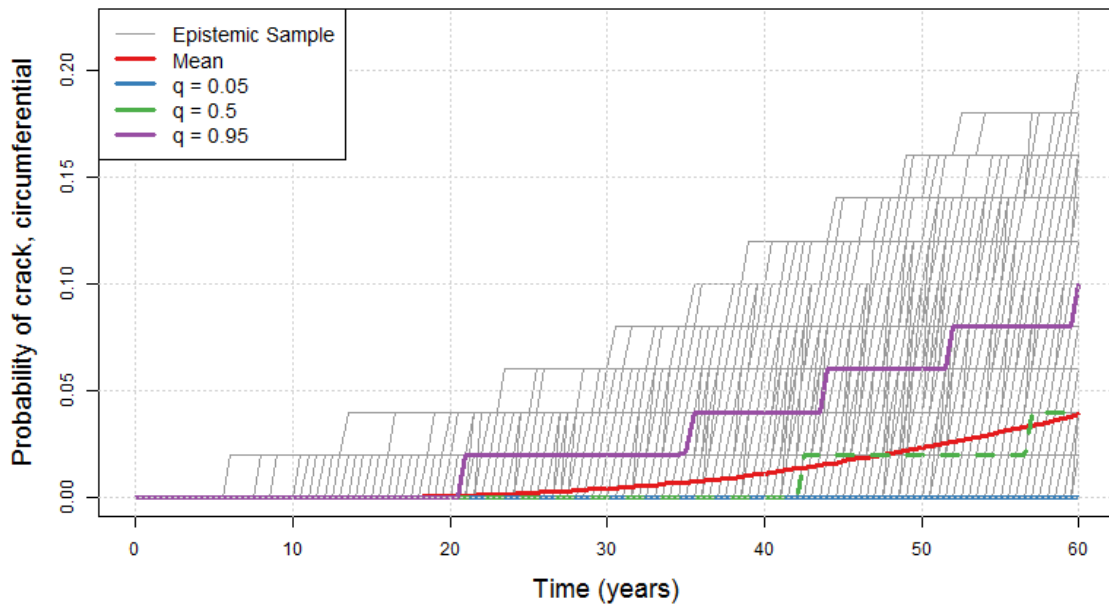


Figure 234: Probability of occurrence of circumferential cracks for each epistemic realization (grey), the mean (red), and the 5th (blue), 50th (green), and 95th (purple) percentiles for Scenario 11, Run 6.

3.11.5 Convergence Analysis

Run 4 (1000/50 LHS IS 0.95) was run with five different random number seeds for both the epistemic and aleatory loops. Figure 235 shows the mean circumferential crack occurrence for each replicate. The replicates produce nearly identical estimates, consistent with solution convergence. The 95% confidence interval is shown in Figure 236. The narrow confidence interval indicates that the 1000/50 LHS IS sampling scheme produces stable and reproducible estimates of the mean probability of circumferential crack occurrence.

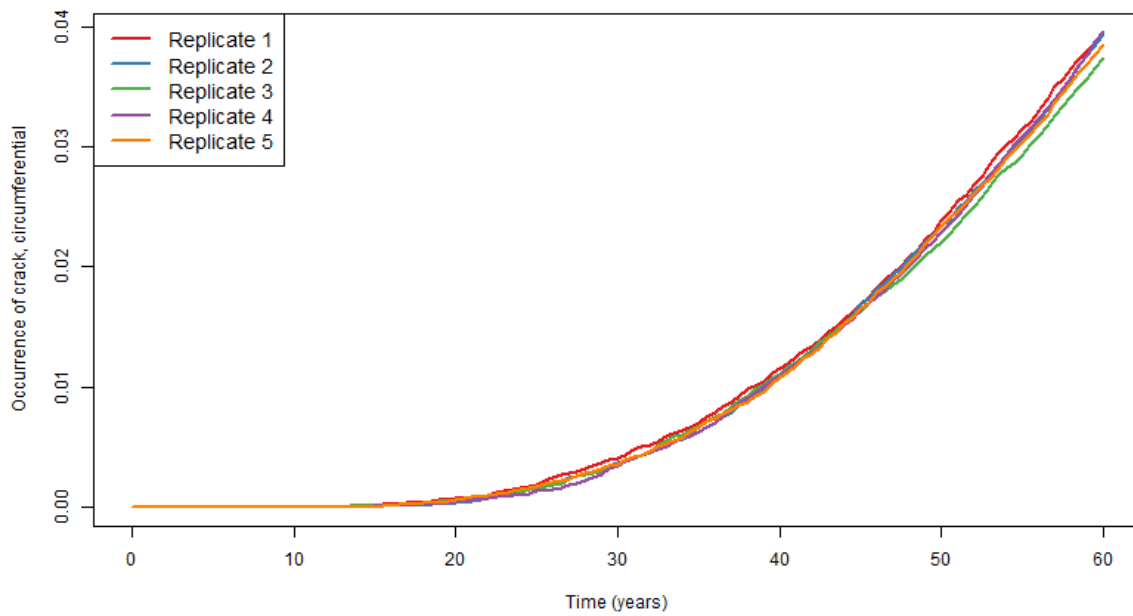


Figure 235: Comparison of estimated mean probability of occurrence of circumferential cracks for five replicates for Scenario 11 Run 4 (1000/50 LHS IS 0.95).

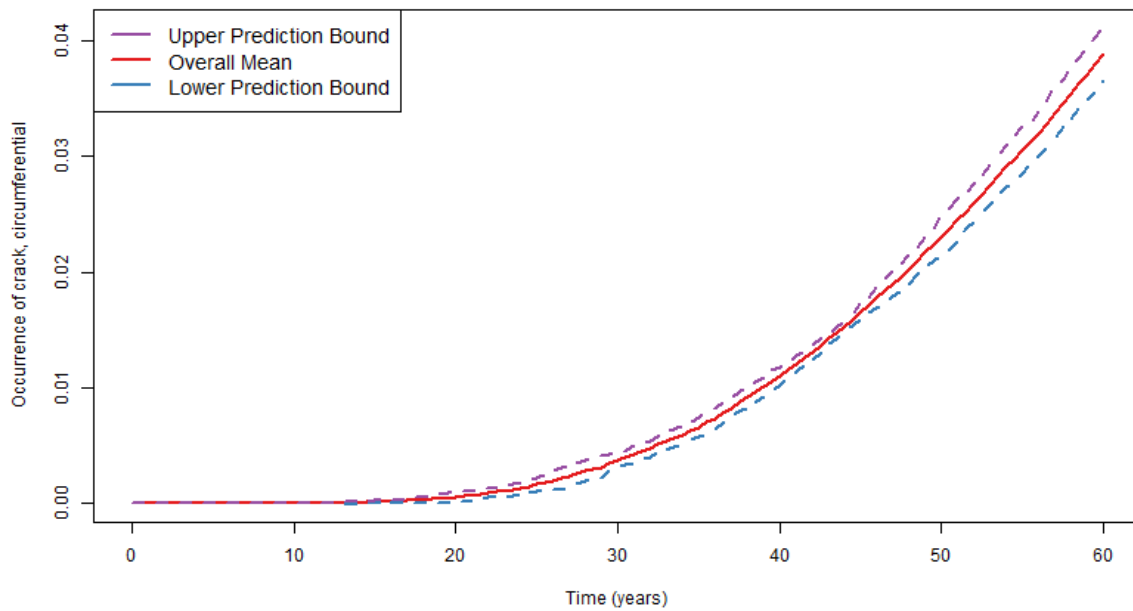


Figure 236: 95% prediction interval over mean probability of circumferential crack for five replicates for Scenario 11 Run 4 (1000/50 LHS IS 0.95).

This page intentionally left blank.

4 ADDITIONAL STUDIES

The broad findings identified across the analysis of the eleven scenarios presented in Section 3 of this report prompted questions that required further investigation. The results of additional studies presented in this section use Scenario 3 as an exemplar. In the following sub-Sections, we explore aleatory sample size selection (Section 4.1), importance sampling quantile selection (Section 4.2), uncertainty characterization under LHS (Section 4.3), SRS dependence on random seed selection (Section 4.4), and bounding low probability events (Section 4.5). In some cases, these case studies represent areas of future work and investigation that are recommended but are beyond the scope of the current report.

4.1 Aleatory Sample Size Selection

In the xLPR Scenario Analysis presented in this report, the number of aleatory samples is fixed to 50. In this section, the impact of aleatory sample size on the estimation of epistemic uncertainty in the probability of occurrence of an event is described.

Three sources of quantifiable uncertainty are addressed in the Scenario analysis for QoIs: 1) epistemic uncertainty, 2) aleatory uncertainty, and 3) sampling uncertainty. Sampling uncertainty stems from finite sampling of both epistemic and aleatory distributions. We have quantified sampling uncertainty from finite epistemic samples in the scenario analysis, but did not account for finite aleatory sample sizes.

4.1.1 Aleatory Sampling Uncertainty in xLPR

For a single epistemic sample, we estimate the output distribution and/or probability of occurrence of an event (leak, rupture, etc.) over time given those 50 aleatory samples. When estimating a probability with only 50 aleatory samples, the probability of occurrence (p) over time is a step function that can only shift in increments of $1/50$; that is, the estimated p over aleatory samples for each time point will be in the set $[0, 0.02, 0.04, 0.06, \dots, 1]$, even though the true p over time is a continuous function.

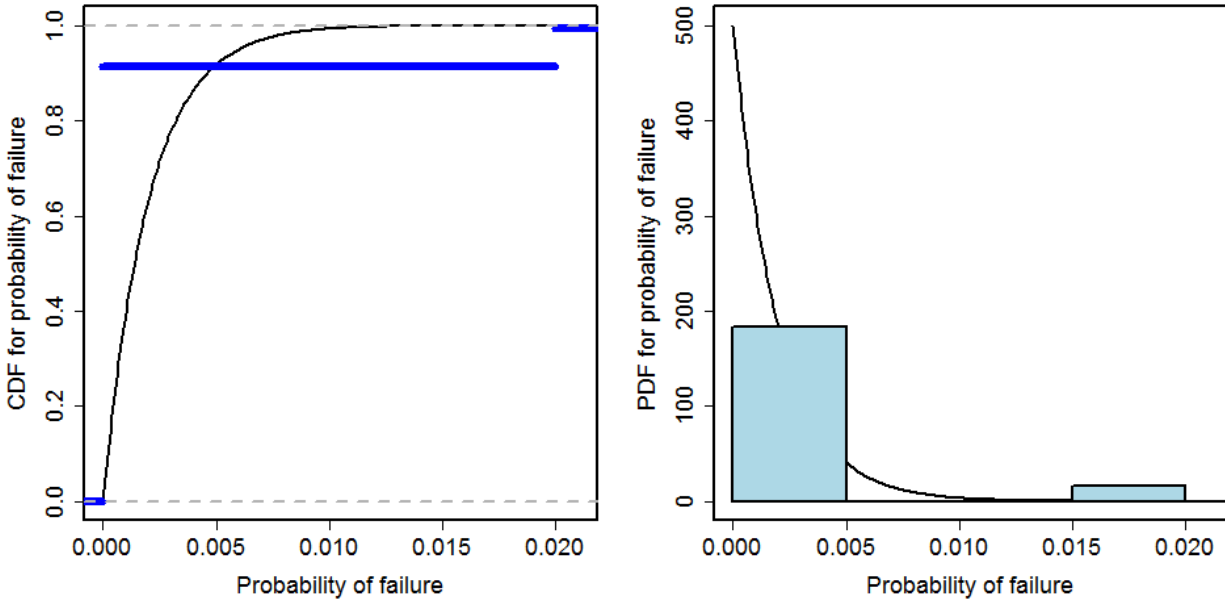


Figure 237: Empirical (blue) and true (black) CDF and PDF over epistemic samples with a small probability of failure; the true estimate assumes no aleatory uncertainty and the empirical estimate is based on $n_a = 50$.

A hypothetical true and empirical CDF and PDF for a low probability of occurrence over epistemic samples with $n_a=50$ are shown in Figure 237. From this figure, it is clear that the estimated probability using $n_a=50$ is not an accurate estimator for the true distribution, and the finite aleatory sample size introduces additional uncertainty into the analysis. When estimating the mean probability of failure over all aleatory and epistemic samples, aleatory sampling uncertainty is likely negligible because we are averaging over n_e different estimates of the same probability (thus reducing aleatory sampling uncertainty). However, when trying to estimate the distribution of probabilities over epistemic samples in order to, for instance, bound epistemic uncertainty on a probability, aleatory sampling uncertainty is non-negligible.

In the scenario analysis described in this report, we calculated epistemic quantiles to characterize the amount of epistemic uncertainty in a QoI, but did not account for aleatory sampling uncertainty. To consider the impact of ignoring aleatory sampling uncertainty in Scenario 11, we increased the number of aleatory samples and compared the results. Specifically, in Figure 238, we estimated the probability of circumferential crack using runs with 100 epistemic and 50 aleatory realizations; and then using 100 epistemic and 1000 aleatory realizations. As expected, when the aleatory sample size is 1,000, the epistemic curves are smoothed out over time, giving a more precise estimate of the probability of circumferential crack over time.

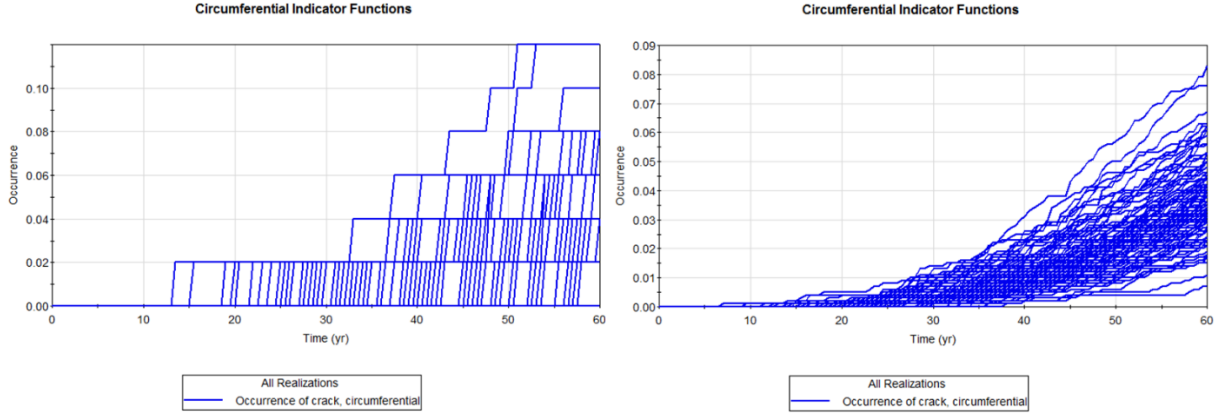


Figure 238: Probability of circumferential crack using 100/50 sampling (left) versus 100/1000 sampling (right). Increasing the number of aleatory samples provides a more precise estimate of the failure probability for each epistemic sample.

4.1.2 Simulated Example

In this section, we provide a simple example illustrating how finite aleatory sample size uncertainty increases uncertainty in the output; ignoring this source of uncertainty can lead to overestimating the amount of uncertainty attributable to the epistemic variables. We consider the time to leak T_L output in the simple model studied by (Jyrkama & Pandey, 2016):

$$T_L = T_I + \frac{W}{R} \quad (14)$$

where T_I is the time to crack initiation, W is the pipe thickness, and R is the crack growth rate.

The objective is to quantify the epistemic uncertainty in the time to leak where we assume W is fixed, T_I is aleatory, and R is the epistemic variable. In other words, we would like to know the amount of epistemic uncertainty attributed the crack growth rate. Though this is a simple model and the categorization of epistemic and aleatory variables is arbitrary, the discussion here applies to more complex models such as xLPR.

We use the same distributions for the variables as (Jyrkama & Pandey, 2016). In particular, T_I is assumed to be Weibull distributed with a shape parameter of 3 and scale parameter of 480 and R is assumed to be normally distributed with mean 5 and standard deviation 1. Finally, $W = 40$ is fixed. The dual loop Monte Carlo applied to this situation first samples the epistemic variable R in the outer loop. Then, conditional on this variable, many samples of the aleatory variable T_I are taken, evaluating T_L each time. The number of samples of the aleatory variables is referred to as the aleatory sample size. The outer loop is repeated many times, referred to as the epistemic sample size. We first concentrate on a single iteration of the outer loop.

Suppose the value of $R = r$ is chosen for the outer loop value. The samples of T_L (one for each aleatory sample of T_I) are samples from the distribution of T_L given $R = r$. The number of aleatory samples in the inner loop determines the precision in which we can estimate this conditional distribution. As an example, we set $R = 5$ and take repeated samples of size 25 (aleatory sample

size). The independent estimates of the conditional distribution of T_L given $R = 5$ are displayed in the top left panel of Figure 239. Each gray curve in the figure is an independent estimate of the cumulative distribution function $P(T_L \leq t | R = 5)$ conditional on $R = 5$ from a sample of T_L of size 25. Note the significant variation around the true value of $P(T_L \leq t | R = 5)$ displayed by the blue curve. This variation can be reduced by increasing the aleatory sample size. For example, the top right, bottom left, and bottom right panels display similar results for aleatory sample sizes of 50, 100, and 1000, respectively. As the aleatory sample size increases, the estimates of the conditional distribution become less variable around the true distribution.

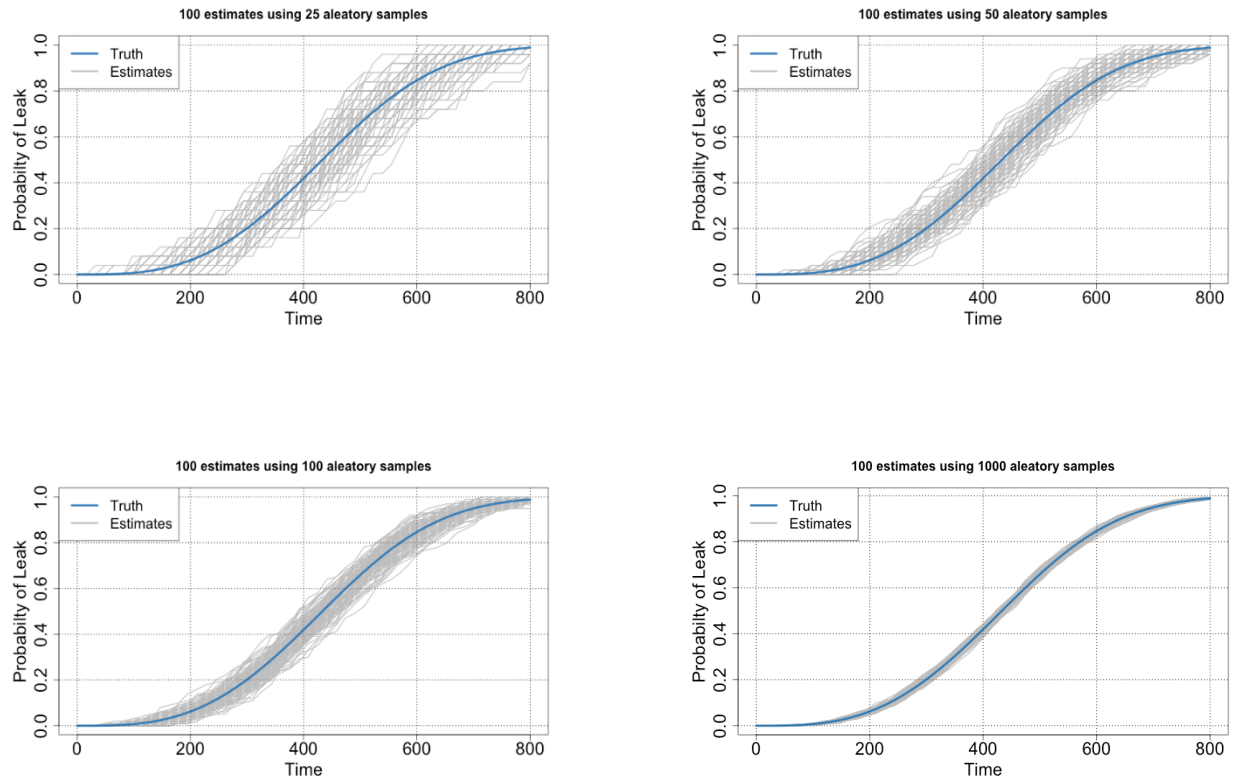


Figure 239: 100 independent estimates (gray) of $P(T_L \leq t | R = 5)$ (blue) for aleatory sample size 25 (top right), 50 (top left), 100 (bottom left), and 1000 (bottom right).

This result has implications for choosing an aleatory sample size. For each outer loop in the Monte Carlo, the goal is to estimate the conditional distribution of the output given the sampled epistemic variable (in this case the conditional distribution of T_L given R). Due to the finite aleatory sample size, there is epistemic sampling uncertainty in this estimate (displayed in Figure 239). This uncertainty propagates through to the analysis of the uncertainty in $P(T_L \leq t | R)$ due to the epistemic uncertainty in R . To see this, consider Figure 240 which displays the dual loop results for 1000 epistemic samples and aleatory sample sizes of 25 (top left), 50 (top right), 100 (bottom left), and 1000 (bottom right). Each gray curve is an estimate of the conditional distribution of T_L given a value of the epistemic variable R (the same set of 1000 epistemic variables is the same for each graph). The variation in the gray curves is due to two sources of uncertainty. One is, of course, the epistemic uncertainty in R . The other source is the sampling uncertainty in each of the conditional

distribution estimates (gray curves) due to the finite aleatory sample size. In this example, we see that as the aleatory sample size grows, the variation across epistemic samples decreases significantly. Thus, the epistemic uncertainty due to R is relatively small compared to the finite aleatory sample size uncertainty. This finite sample uncertainty is important to understand when the objective is to quantify the output uncertainty due to the epistemic uncertainty in R , i.e., when trying to estimate epistemic uncertainty intervals. Under an insufficient aleatory sample size, we risk attributing uncertainty to epistemic variable R that is due to finite aleatory sample size uncertainty.

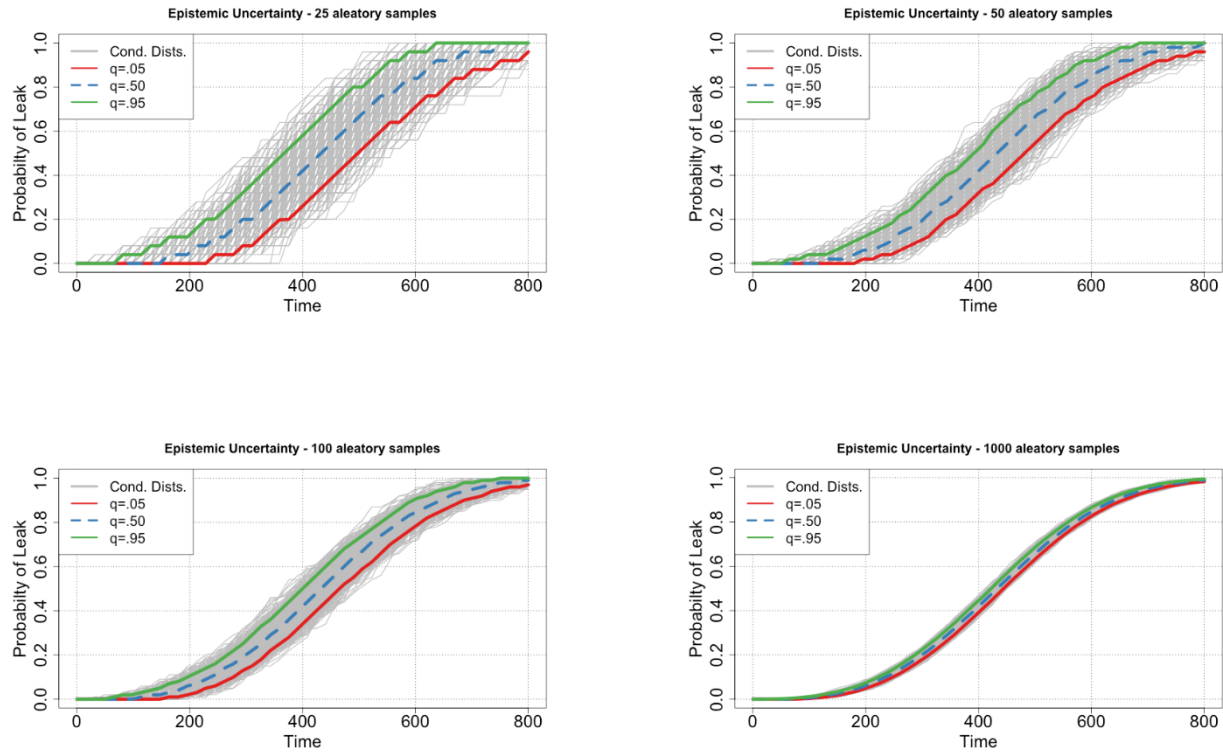


Figure 240: Figure 3 Dual-loop results for 1000 epistemic samples and aleatory sample sizes of 25 (top left), 50 (top right), 100 (bottom left), and 1000 (bottom right). The blue dashed line is the median of these estimates across the epistemic samples and the green and red lines are the 0.95 and 0.05 quantiles.

4.1.3 Aleatory Sample Size Selection Summary

The following list summarizes the results of the aleatory sample size study described in this section:

1. Three sources of uncertainty exist in an output of interest from a dual-loop Monte Carlo analysis:
 - a. Uncertainty due to the epistemic variables.
 - b. Uncertainty due to the aleatory variables.
 - c. Sampling uncertainty due to finite sampling of both aleatory and epistemic distributions
2. In the methodology presented for each scenario in this report, sampling uncertainty in the aleatory variables and epistemic uncertainty are not separated when analyzing distributions across epistemic samples. More than 50 aleatory samples are needed to reduce the amount of uncertainty due to finite aleatory sample size to get good estimates of the uncertainty due to the epistemic variables; on the other hand, an aleatory sample size of 50 is likely sufficient to estimate the mean probability, rather than epistemic uncertainty intervals.
3. The number of aleatory samples needed is dependent on the relative uncertainty in the output due to the aleatory and epistemic variables. In the above example, most of the uncertainty in T_L is due to T_I . Suppose instead T_I was considered epistemic and R considered aleatory. Then the variation in T_L across R for a fixed value of T_I is relatively small and fewer aleatory samples would be needed to obtain a sufficient estimate of each conditional distribution $P(T_L \leq t | T_I)$.

4.2 Importance Sampling Quantile and Methods

Under the methods used to analyze each scenario in this report and with the inputs currently defined for each scenario, we did not see any clear gains in efficiency using importance sampling when estimating the mean probability of circumferential rupture. Importance sampling should be beneficial for estimating rare events and when the importance sampling distribution is selected to target variance reduction in the QoI. Specifically, targeting the input distribution tails (for sensitive variables) using IS will help increase the number of event occurrences. When the probability of occurrence is rare, substantial efficiency can be gained using IS. When the probability is not rare, IS sampling will target a relatively low-importance region of the input space at a very high rate, decreasing efficiency.

Poor selection of the importance distribution can substantially reduce efficiency in QoI estimates, so selection of the importance distribution should be handled with care. For a single IS input, the default IS distribution in GoldSim is a mixture of the nominal distribution and a truncated normal distribution centered at a target quantile (we only considered IS for one variable based on the sensitivity analysis results). The truncated normal distribution is defined in quantile space. Specifically, the user selects the target quantile distribution mean and, when only one variable is selected for IS, GoldSim defaults the range to be a 1% quantile width around the target and defaults the standard deviation to be 0.25%. We selected the 0.95 quantile as the target for the DM1 multiplier, which successfully increased the number of overall events (cracks, ruptures, and leaks). The efficiency of IS will likely change depending on the selected form of the importance sampling distribution as well as the targeted quantile (when using mixture importance distributions as in GoldSim).

4.2.1 Simulated Example

We use a simple example to illustrate how efficiency gains in importance sampling change with the importance distribution (including both the shape and target quantile) and the probability of occurrence. We simulated an input $X \sim N(\mu, 1)$ and calculated $p = \Phi(X)$, where Φ is the CDF of the $N(\mu, 1)$ normal distribution. Hence, p is between 0 and 1 and can be interpreted as a probability. The probability becomes closer to 0 as μ decreases. We can consider importance sampling X to improve estimation of the mean of p , denoted \bar{p} using Monte Carlo simulation. This example is similar to the setup of xLPR, but there is only one input variable and we ignore the inner aleatory loop for convenience and assume that p is known exactly.

We consider two importance distributions: (1) a mixture of the nominal distribution $N(\mu, 1)$ and a uniform surrounding a target quantile and (2) a mixture of the nominal distribution $N(\mu, 1)$ and a truncated normal surrounding a target quantile. For the uniform, we use a range of [(target - 0.005) quantile, min(1-1e-10, target + 0.005) quantile]. For the truncated normal, we sample on the quantile scale and then transform to the original input space. We use standard deviation of 0.0025 and a range of [target - 0.005, min(1-1e-10, target + 0.005)]. We consider $\mu = [-1, -3, -5, -6, -7]$ and target quantiles [0.95, 0.995, 0.999]. By varying μ , the magnitude of probability p then varies between 10^{-7} and 10^{-1} .

We estimate the mean of p by taking 1000 Monte Carlo simulations from X , calculating $p = \Phi(X)$, and then calculating \bar{p} using the empirical mean over all 1000 samples. We repeat this process 1,000 times and calculate the standard deviation of the 1,000 different estimates of \bar{p} to characterize

sampling uncertainty. We conduct this Monte Carlo simulation using both simple random sampling and importance sampling (considering the two different densities) and quantify efficiency gains using IS based on the ratio of the standard deviation of \bar{p} under SRS and the standard deviation of \bar{p} using IS.

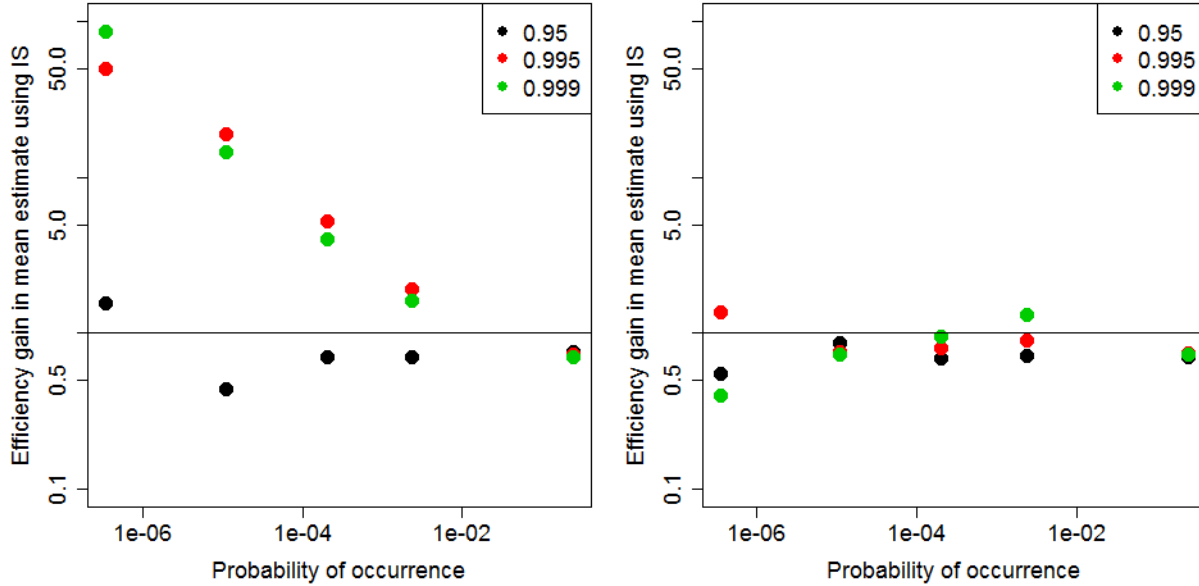


Figure 241: Efficiency gains using IS as a function of the true probability of occurrence and the targeted IS quantile for the (left) normal-uniform mixture importance distribution and (right) the normal-truncated normal mixture distribution. Both the x- and y-axes are on the log scale. A horizontal line is drawn at 1, where there is no difference in efficiency between SRS and IS. The probability of occurrence is simply a function of μ .

Results are shown in Figure 241. The gains using IS with the normal-uniform distribution increase markedly as p nears 0 when the 0.995 or 0.999 quantiles are targeted. When the probability of occurrence is not rare, sampling these extreme quantiles can decrease efficiency, because the input distribution tails are heavily oversampled in this case. Targeting the 95th percentile is less efficient than SRS in this example. Using the normal-truncated normal importance distribution, IS does not provide improvements in efficiency over SRS.

4.2.2 Scenario 3 Importance Sampling Quantile Study

To assess variability in results according the importance distribution target quantile, we conducted additional runs for Scenario 3. To gain efficiency in estimating rare probabilities, the previous Importance Sampling Width Function Study (xLPR-SDD-FW, 2016) recommends targeting the 0.995 quantile. Therefore, we conducted additional Scenario 3 runs using the 0.995 quantile to compare to the 0.95 quantile. Further, a general rule in importance sampling is that the importance distribution should have tails at least as heavy as the nominal distribution. The default importance distribution in GoldSim (nominal-truncated normal mixture) does not necessarily meet this criterion. However, when the target quantile is greater than 0.995, the importance distribution changes to a mixture of the nominal distribution and a uniform distribution on the upper tail. Thus, we conducted additional Scenario 3 runs using the 0.999 quantile for comparison.

We considered four different sampling schemes using SRS 100/50 and LHS 1000/50 samples:

1. No IS
2. IS at 0.95 target quantile
3. IS at 0.995 target quantile
4. IS at 0.999 target quantile

We also considered the no IS and .95 target quantile runs using 500/50 samples with both SRS and LHS. We added one additional run at the .999 quantile using SRS 1000/50.

We compare three metrics:

1. Time at first rupture, circumferential
2. Estimated mean probability of rupture, circumferential, up to 30 years
3. Relative confidence interval width for mean probability of rupture, circumferential, up to 30 years.

4.2.2.1 Time at first circumferential rupture

The times of first circumferential rupture in years are shown in Table 13 below for the different sampling schemes that were analyzed.

Table 13: Time (yr) of first circumferential rupture for various sampling schemes.

	No IS	IS, 0.95	IS, 0.995	IS, 0.999
SRS, 100/50	19.5 yrs	15	12	12
SRS, 500/50	12	6.5	Not Run	Not Run
SRS, 1000/50	8.5	5.5	Not Run	4.5
LHS, 500/50	4.5	5	Not Run	Not Run
LHS 1000/50	7.5	7	6	6

Using SRS with IS decreases the time of first rupture. For the 100/50 and 1000/50 runs, the time of first rupture decreases as the quantile increases, as expected. Using LHS, the time of first rupture is similar for no IS and IS at the .95 quantile; for the 1000/50 runs, the time of first rupture decreases as the target quantile increases.

4.2.2.2 Mean probability of circumferential rupture up to 30 years

SRS. In Figure 242 and Figure 243, we see that, under SRS, importance sampling generates cracks earlier and thus allows estimation of the probability of rupture at early time points. We ran the SRS 100/50 IS runs with 2 random seeds. There is a substantial amount of variability in the estimated probability of rupture between the two random seeds, suggesting we should not generalize results from these un-converged runs. For both seeds, IS using the .995 and .999 quantiles gave similar results. We ran an additional SRS 1000/50 run using IS at target quantile .999. Targeting the .999 quantile using IS seems to provide a much smoother estimate of the probability of rupture over time, suggesting better convergence using IS. We did not run the .995 quantile for n=1000.

LHS. In Figure 244, we see that, under LHS, importance sampling generates cracks earlier when $n=1000$, but not when $n=500$. When $n=1000$, IS using the .995 and .999 quantiles produced very similar results (we could not make a comparison for smaller sample sizes because these runs were not conducted); targeting these quantiles, the estimated probability of rupture is somewhat smoother at earlier time points (between years 7-12).

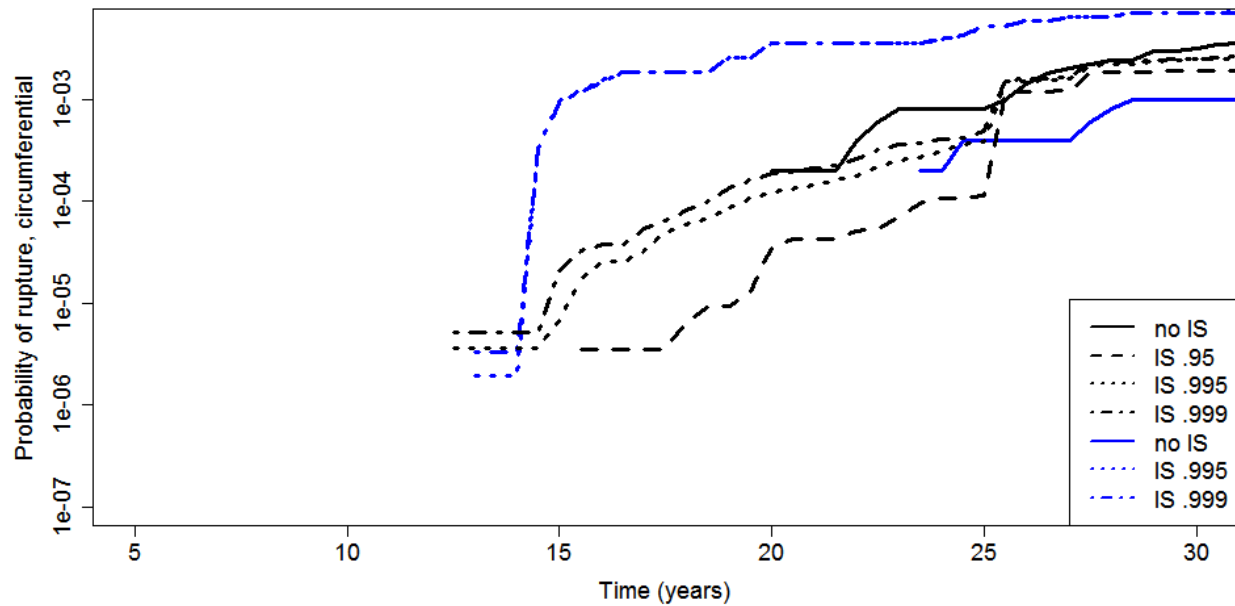


Figure 242: Probability of circumferential rupture on log scale from time 5 years to 30 years across different SRS sampling schemes with 100 epistemic samples. Black lines correspond to the original random seed and blue lines are a different randomly selected random seed.

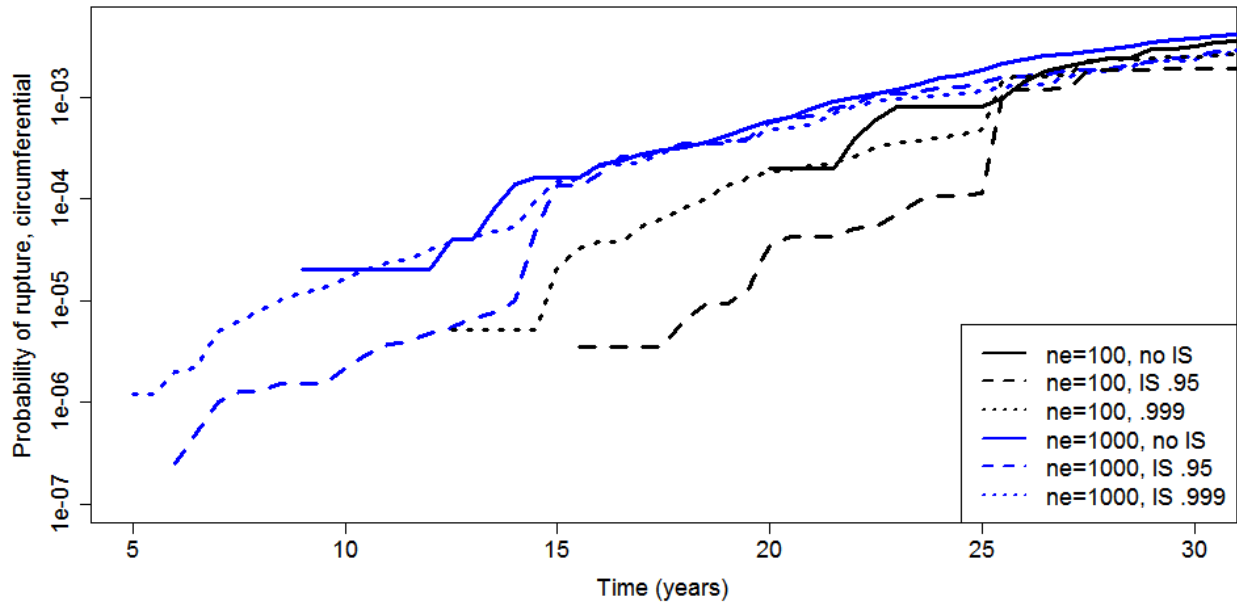


Figure 243: Probability of circumferential rupture on log scale from time 5 years to 30 years across different SRS sampling schemes with different numbers of epistemic samples.

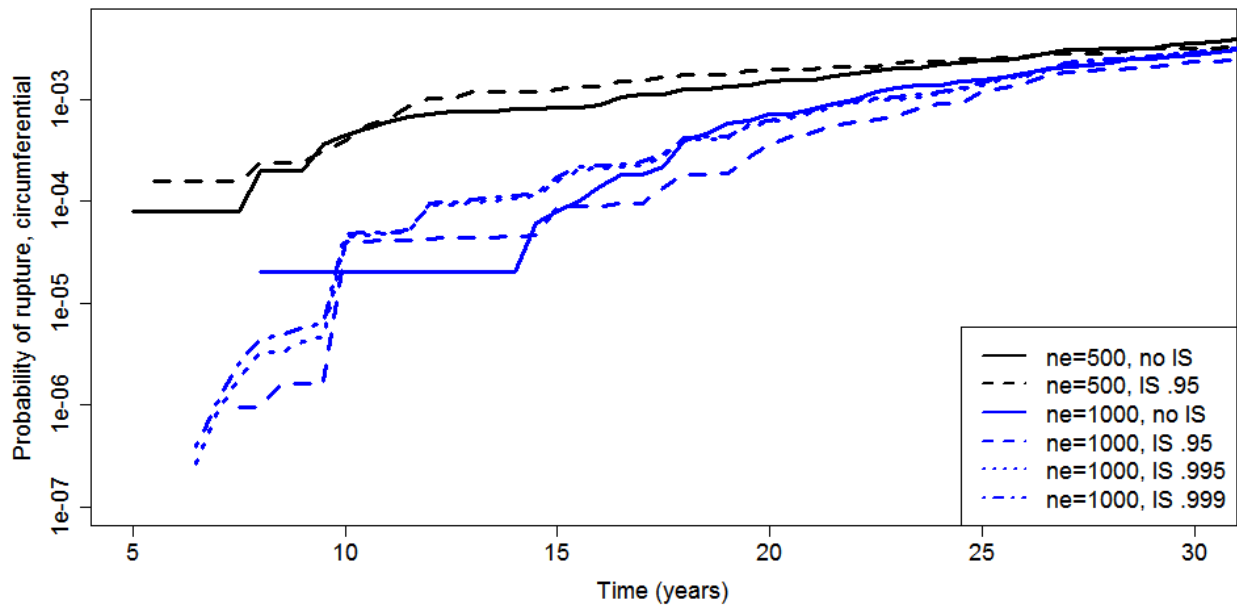


Figure 244: Figure 3: Probability of circumferential rupture on log scale from time 5 years to 30 years across different LHS sampling schemes.

4.2.2.3 Relative confidence interval width

We compare relative confidence interval width across sampling scenarios, where relative CI width is defined as the confidence interval width divided by the estimated mean probability. We use relative confidence interval width, rather than actual confidence interval width, because confidence interval width is highly correlated with the estimate (e.g. probabilities on the order of 10^{-3} will also have confidence interval widths on the order of 10^{-3}). So, when comparing sampling schemes, it is simpler to interpret confidence interval width relative to the actual estimate.

SRS. In Figure 245, we see that under SRS 100/50 sampling, there are not any clear trends in confidence interval width. For one seed, IS gives smaller relative CI width, while for the other seed, the relative CI width is similar. In Figure 246, with $n=1000$, we do not see a clear improvement in relative CI width for the IS runs.

LHS. In Figure 247, we still do not see a clear difference in relative CI width using IS versus SRS, regardless of the targeted quantile. IS targeting the .995 and .999 quantiles does have lower relative CI width at early years (7-15 years).

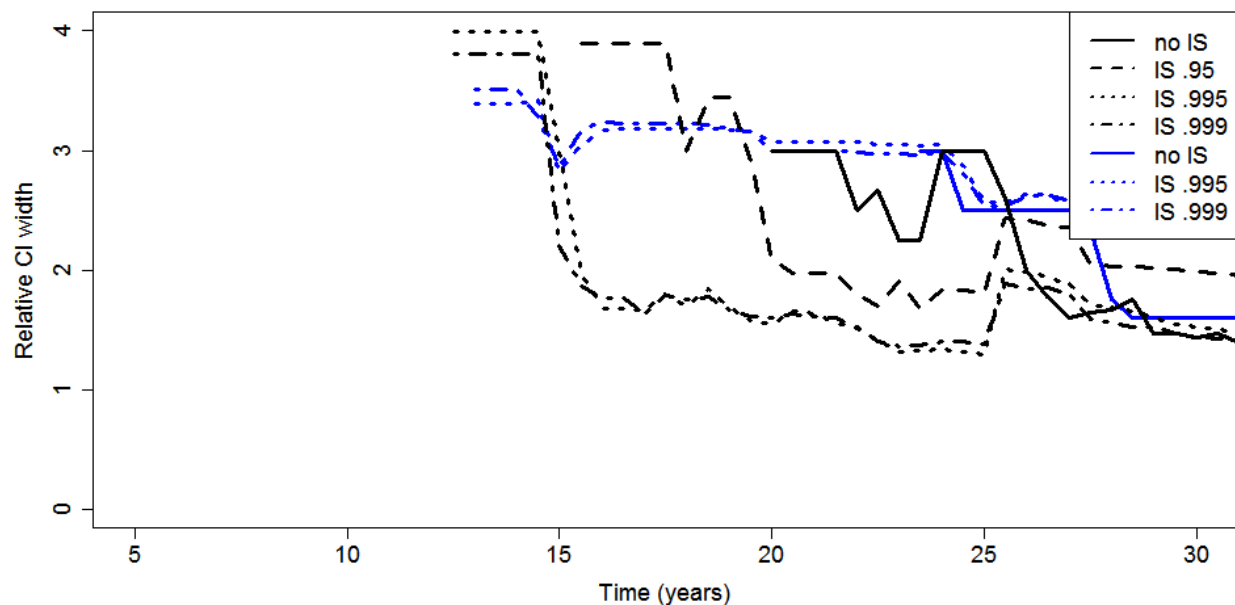


Figure 245: Bootstrap relative confidence interval width for probability of circumferential rupture on log scale from time 5 years to 30 years across different SRS sampling schemes with 100 epistemic samples.

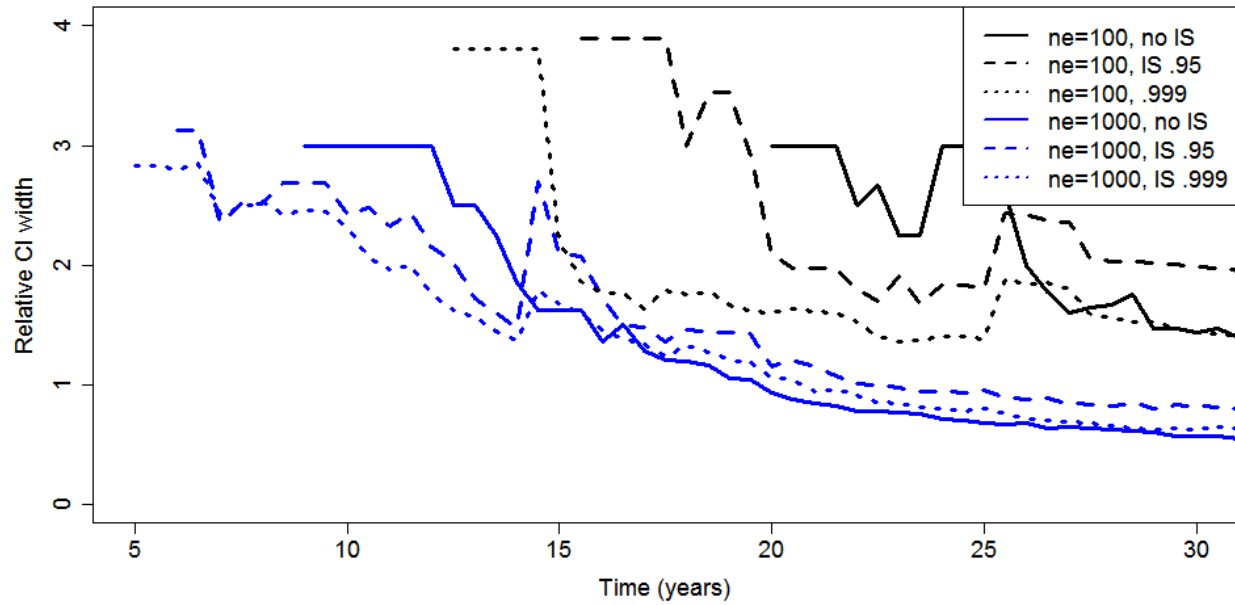


Figure 246: Bootstrap relative confidence interval width for probability of circumferential rupture on log scale from time 5 years to 30 years across different SRS sampling schemes with different numbers of epistemic samples. We omitted results for $n = 500$ as they were similar to the other sample sizes.

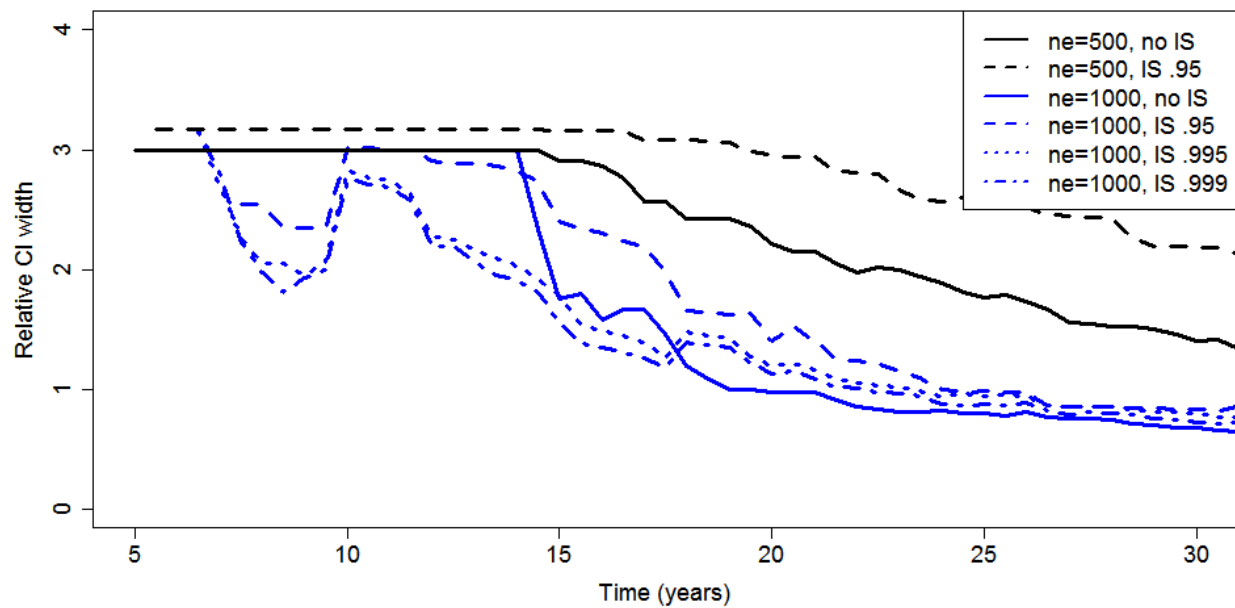


Figure 247: Bootstrap relative confidence interval width for probability of circumferential rupture on log scale from time 5 years to 30 years, across different LHS sampling schemes. Confidence intervals are calculated treating the data like an SRS and therefore are not accurate, but relative widths between schemes should be comparable.

4.2.2.4 Summary

We cannot draw many generalizable inferences about the probability of rupture using IS versus SRS. In our examples, using IS with SRS sampling does decrease the time of first rupture. Using LHS also decreases the time of first rupture relative to SRS. Using IS with LHS did not have a marked impact on time of first rupture relative to LHS without IS. IS using the .995 and .999 quantiles produced similar estimates of the probability of rupture in our examples. There were not any clear trends regarding relative confidence interval widths.

As a final example, we plot the estimated probability of rupture and confidence bounds for the SRS 1000/50 runs with no IS and IS targeting the .999 quantile. Results are very similar between the two runs, aside from the fact that IS allows earlier estimation of the probability of rupture (IS does provide a non-zero lower bound on the estimated mean probability of rupture up to 15 years; however, it is not clear if estimating a lower bound is substantively interesting in practice).

From this study, we conclude that more runs would be needed to draw generalizable inferences about how to effectively use IS in the scenario analysis.

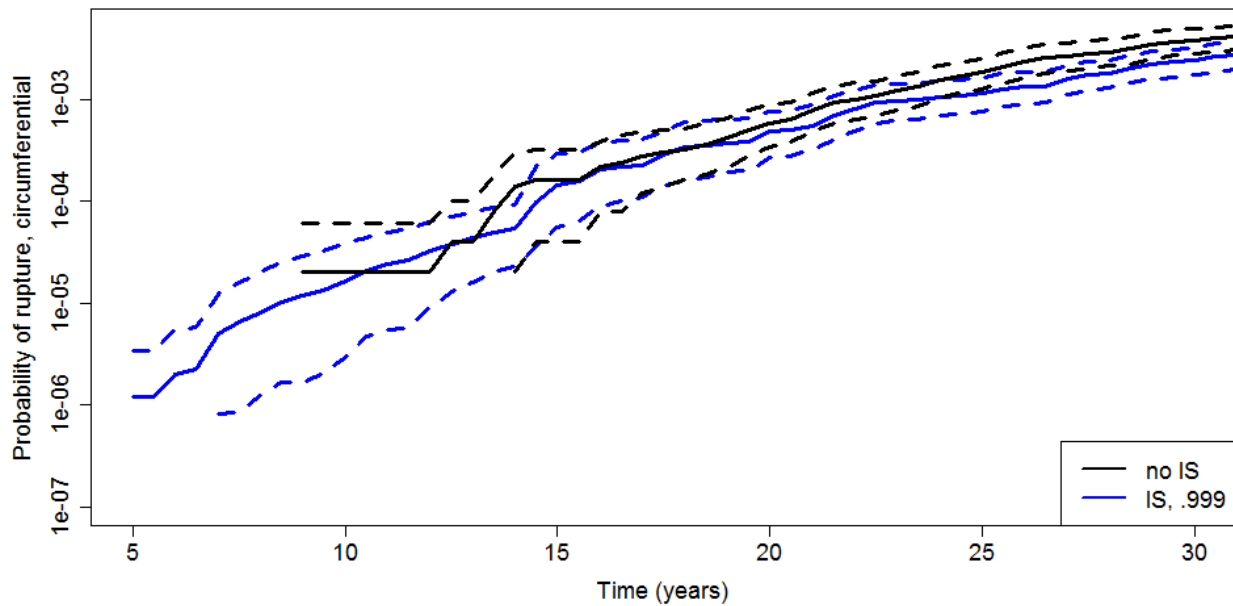


Figure 248: SRS 1000/50 runs using no IS versus IS targeting the 0.999 quantile. IS seems to improve inferences from years 5-15 where IS seems to give a smoother and more precise estimate of the failure probability.

4.2.3 Importance Sampling Quantile and Methods Summary

More investigations are needed regarding optimal selection of importance distributions for xLPR applications before providing generalizable recommendations pertaining to how to importance sample for estimating rare probabilities. It is noteworthy to mention that such investigations will depend on the QoI. Selection of the target quantile and the importance distribution will impact the

efficiency with which rare event probabilities are estimated. Importance sampling distributions should be selected with care.

4.3 Sampling Uncertainty Using Latin Hypercube Sampling

In the scenario analysis described in this report, we do not calculate sampling uncertainty under the LHS design. Rather, we use SRS sampling uncertainties as a conservative upper bound for the LHS case. There is not a simple analytic form for an unbiased estimate of sampling uncertainties using LHS (Stein, 1987; Helton & Davis, Latin Hypercube Sampling and the propagation of uncertainty in analyses of computer systems, 2003) and bootstrapping an LHS sample is not possible. In general, LHS should be at least as efficient as SRS for the xLPR scenario analysis, though it is not clear how much efficiency is gained using LHS.

We summarize some results pertaining to the efficiency of LHS relative to SRS. The EPA's "Guiding principles for Monte Carlo analysis" recommends using LHS over SRS with complex models or when time and resource constraints are present (Firestone, et al., 1997). (McKay, Beckman, & Conover, 1979) give analytic results showing the efficiency of LHS relative to SRS for estimating means, variances, and CCDFs. (Stein, 1987) illustrates that LHS is always more efficient than SRS and shows that the efficiency of LHS increases with the additivity of the output as a function of the inputs. (Manteufel, 2000) provides numerical study results comparing LHS to SRS for estimating the mean, variance, and a tail percentile and concludes that LHS never performs worse than SRS. LHS is shown to have the greatest efficiency gains with simpler models (inputs are bounded and output is a linear function of inputs); in this case, LHS can improve efficiency in mean estimation by two orders of magnitude. On the other hand, LHS is shown to have less improvement over SRS with complex models (infinite tailed inputs with multiplicative interactions on the output).

In the sensitivity analyses results for the scenarios described in this report, there are a large number of uncertain inputs, but only a small set of highly sensitive inputs (e.g., p2543). Further, the output is generally expected to be a rather non-linear function of the inputs. Therefore, it is not clear how much is gained by using LHS in the scenario analysis. To compare the efficiency of LHS and SRS in the scenario analysis, we evaluated a test case, calculating the variability in QoI estimates using 10 LHS runs at different random seeds; for each run, we sampled 100 epistemic and 50 aleatory samples using LHS. We compared the variability in the 10 LHS runs to an analogous 10 runs using SRS by estimating the empirical standard deviation of the mean probability of circumferential rupture across the 10 runs. We found that the estimated variability in the mean probability of circumferential rupture was lower using LHS, as expected (see Figure 249 and Figure 250). In the sample runs, the standard deviation decreased by about a factor of 2 using LHS as compared to SRS. The small number of runs and single test case limits the generalizability of the results to other cases, but the findings do suggest that LHS is preferable to SRS for decreasing sampling uncertainty. Advantages of SRS include simplicity in combining results across different runs as well as simplicity in calculating sampling uncertainty.

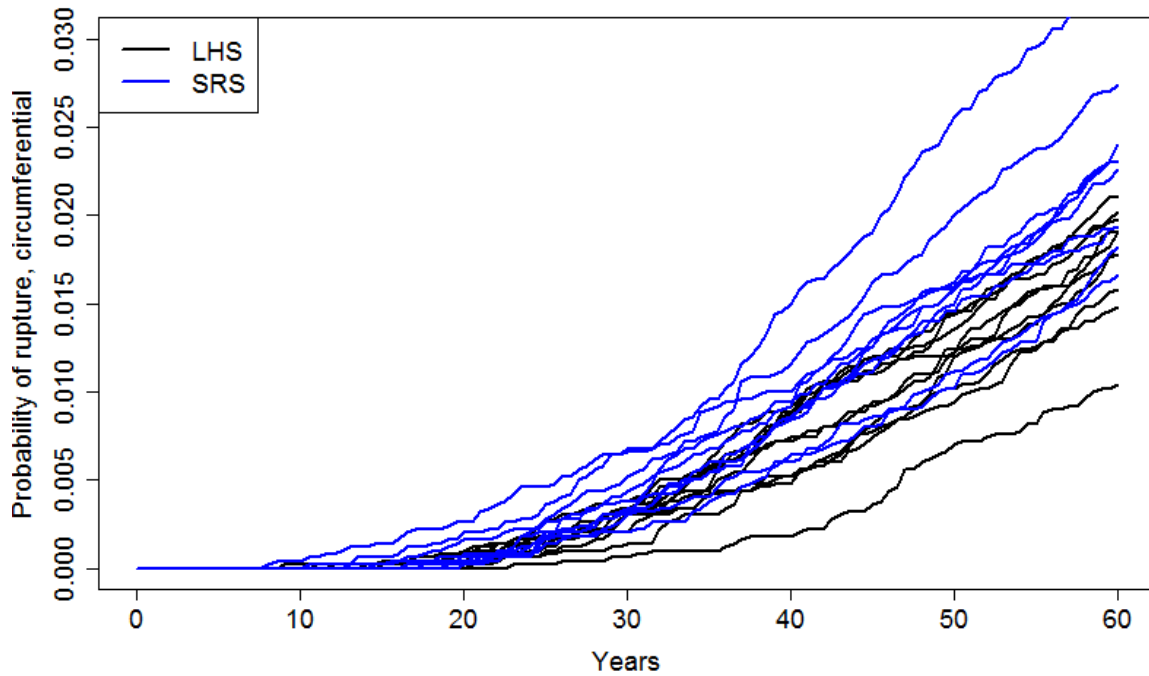


Figure 249: Comparison of 10 different 100/50 epistemic/aleatory realizations of the probability of circumferential rupture using LHS and SRS sampling.

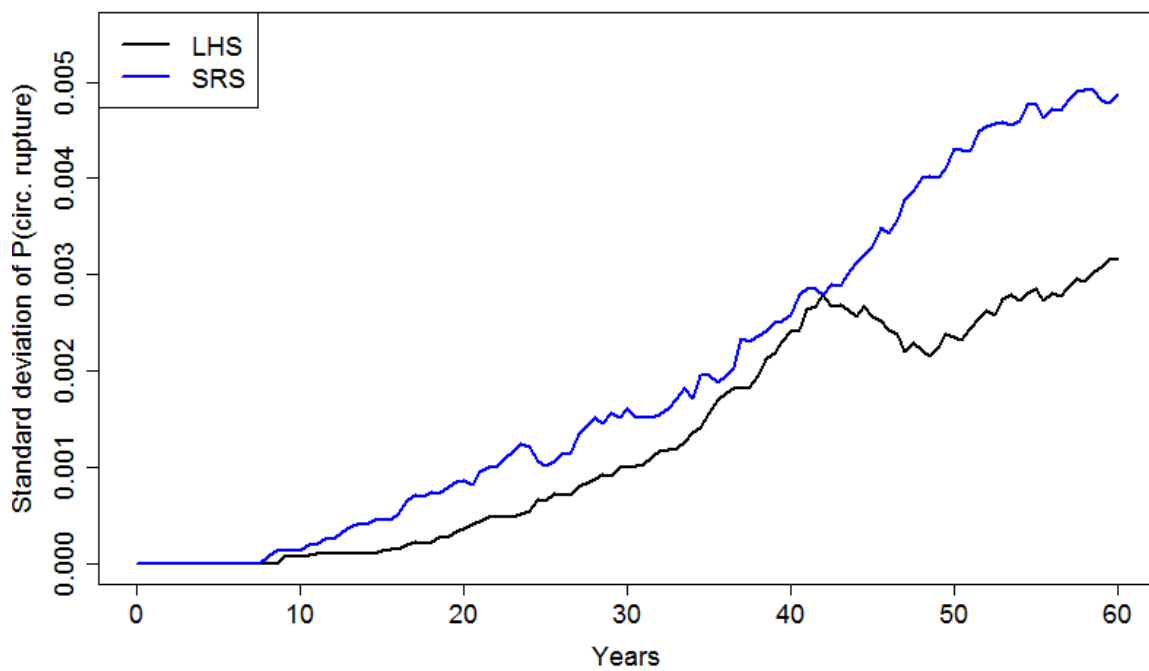


Figure 250: Estimated standard deviation of the probability of circumferential rupture across the 10 different realizations for the LHS (blue) and SRS (blue) samples.

4.4 SRS Dependence on Random Seed Selection

The analysis of the various sampling schemes used for Scenario 3 found that sampling schemes using SRS with no IS, irrespective of sample size, behave differently than the rest of the runs, as demonstrated in Figure 251. In particular, the figure shows the three SRS-only runs have a mean probability of circumferential crack larger than the others; and this occurs regardless of sample size. This result seemed suspect so a separate convergence study was undertaken to examine this result.

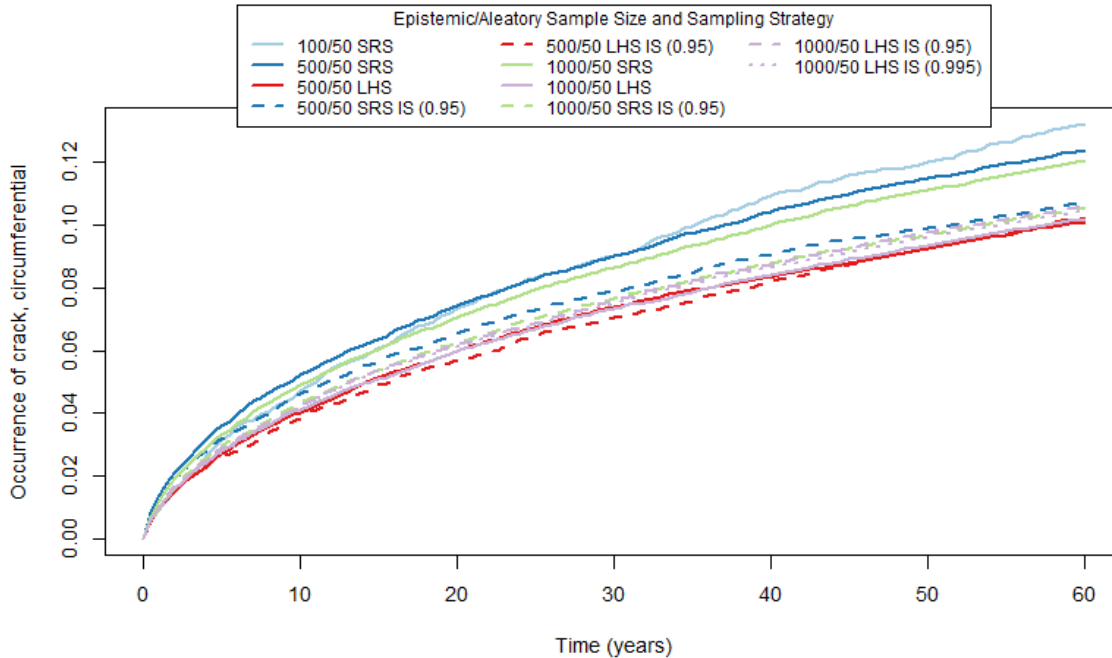


Figure 251: Mean probability of occurrence of circumferential cracks for Scenario 3, Runs 1 through 10.

The first part of this convergence study on SRS-only runs used four replicate Scenario 3 simulations with 500 epistemic, 50 aleatory realizations using SRS without IS with different random seeds. The random seeds for these four replicates were selected by incrementing both the epistemic and aleatory random seeds by 1 for each subsequent run from the original epistemic random seed of 1 and aleatory random seed of 5. The results of this study are shown in blue in Figure 253. The mean of each replicate was found to be greater than the mean for all of the other runs using IS and/or LHS for Scenario 3. This result did not provide any additional insight into the divergence of the mean SRS results from the rest of the Scenario 3 sampling schemes.

The second part of the convergence study on SRS-only runs used an analysis of the sampled values for variable p2543 (the most important variable for this scenario) with many different random seed combinations. This study was computationally much quicker than analyzing the outputs of full replicate simulations as only the sampling portion of the xLPR code was required; all of the other code calculations were removed by deleting the 'xLPR_Model' container at the model root and all related dependencies so that a greater number of results could be obtained more quickly. This did not affect the sampling calculations as all modifications were made downstream of value selection.

SRS samples of size 1000 were collected using 30 different random seeds. The distribution of the input variables is constructed by transforming uniform random samples, so the distribution of the uniform random samples was analyzed for simplicity (p2543 is a highly skewed lognormal variable, and so estimated means and standard deviations are highly variable). The median of the uniform random variable was calculated across the 1000 samples and compared the observed medians to the expected median for data drawn from the same uniform random variable. The results of this comparison are shown in Figure 252.

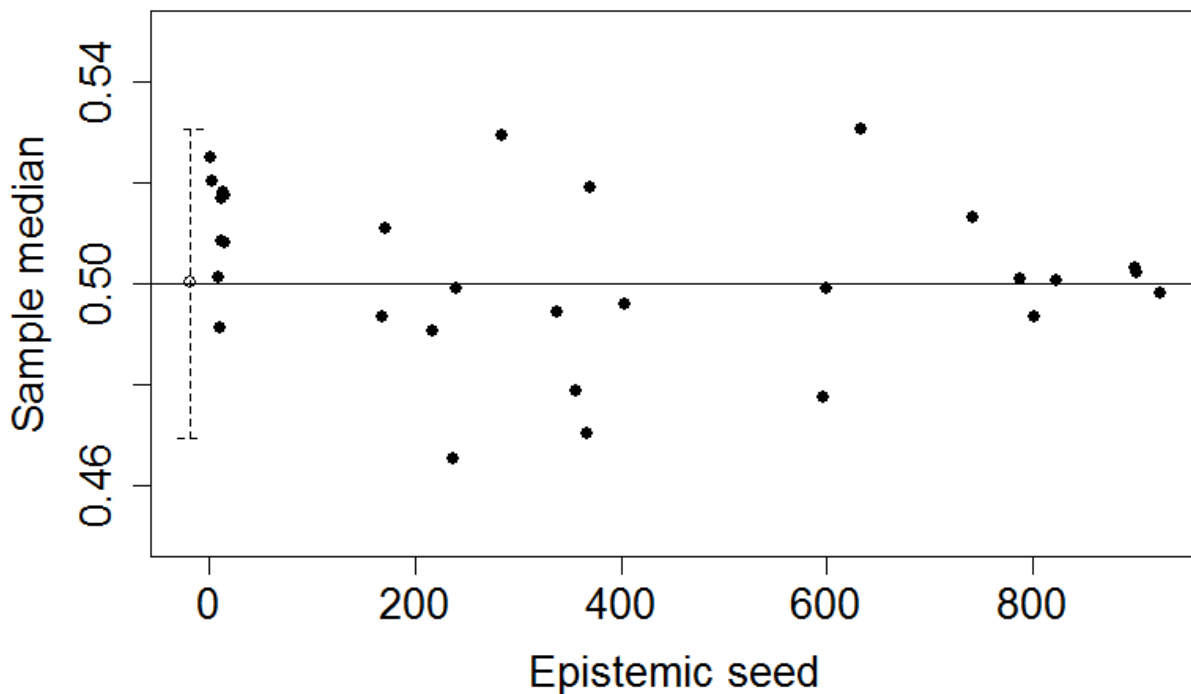


Figure 252: Median of the sampled uniform random variable from 1000 samples as a function of the epistemic seed (black dots). The white dot and dashed interval reflect 95% confidence intervals on the median from sample of size 1000.

It should be noted that, at first, random seed pairs were selected in a pattern of increasing the random seed by 1 for each replicate, as was described above. When these lower random seeds yielded relatively many sampled distributions of p2543 which were higher on average, random seeds from 0 to 1000 were externally selected randomly for both the aleatory and epistemic loops and analyzed. As shown in Figure 252, most of the first set of lower-valued and ordered epistemic random seeds (including the random seeds used for all of the Scenario 3 runs) have higher sampled values on average of variable p2543 when compared to the expected values found in the distribution for p2543 selected in the input set. One of the random seeds with a median lower than 0.5 (epistemic random seed of 10) was used to run a full Scenario 3 simulation, the results of which are shown in green in Figure 253 below.

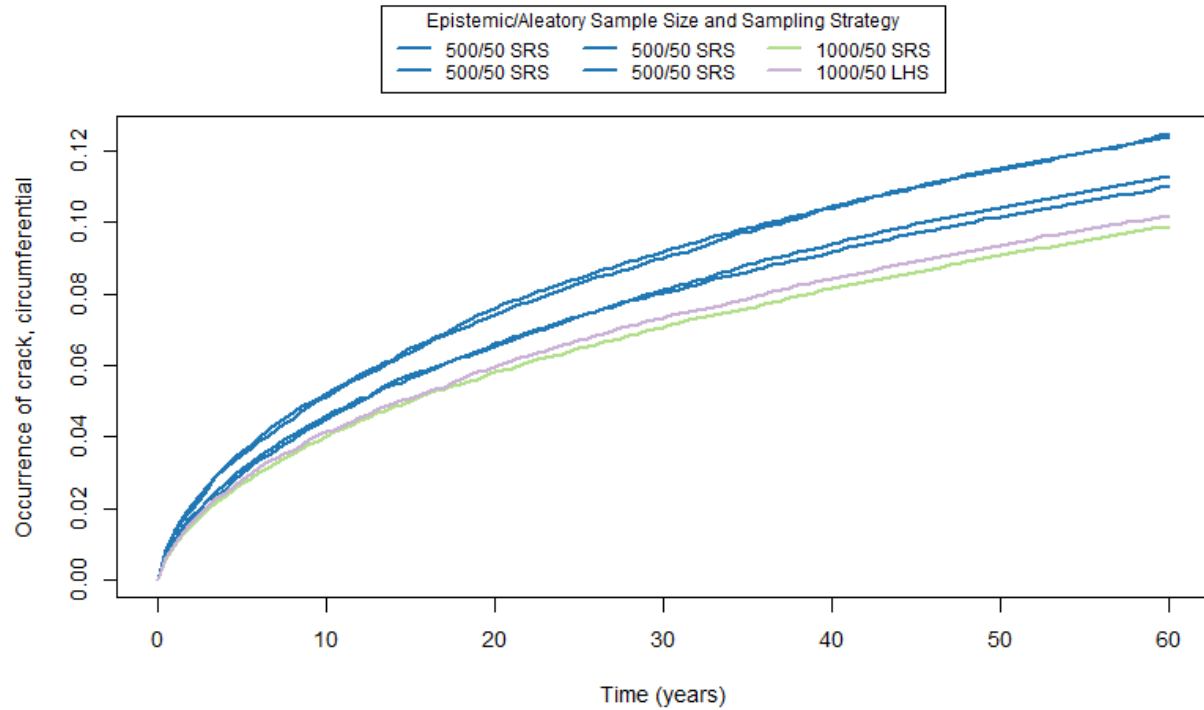


Figure 253: Mean probability of occurrence of circumferential cracks for Scenario 3 runs of 500 epistemic, 50 aleatory realizations using SRS (blue) and 1000 epistemic, 50 aleatory realizations using SRS (green) and LHS (purple) with varying random seeds showing the dependence of SRS results on the selected random seed.

The mean outputs for this simulation were found to be lower than those for the converged sampling schemes using IS and/or LHS, as expected based on the correlation of p2543 with the outputs of interest for this scenario. Hence, there does not appear to be any systematic issues with using SRS, but the random seed should itself be randomized to avoid any potential correlations in sampled values between neighboring seeds.

4.5 Bounding Low Probability Events

As one of the main purposes of xLPR is to reliably estimate very small failure probabilities, it is reasonable to assume there will be cases where it is difficult to observe failures due to computational time limitations. If after conducting a large number of runs zero failures are observed, it may be of interest to give an overall bound on the probability of failure based on the number of simulations that were run.

In the Scenario Analysis effort, we did not see occurrences of rupture with leak rate detection (LRD) in any of the runs over all 11 scenarios. We ran two additional sets of runs. The first run used 10 sets of simulations with different, randomly-selected epistemic and aleatory random seeds and 1000 epistemic, 100 aleatory realizations. The results of these simulations were combined to create a large simulation set with 10,000 epistemic and 100 aleatory realizations. The second run used a modified input set with all samples set to aleatory and combined sets of 1000 epistemic, 100 aleatory realizations to create a set of $1e6$ realizations independent of the dual sampling loop structure. The method for obtaining this sample size is described in detail in Section 4.5.3. In both of these runs, we also did not see any occurrences of rupture with LRD. We now consider the question of bounding the failure probability given these two sets of runs; specifically, we address the question ‘what upper bound can we place on the probability of failure given these runs?’ Conclusions made here apply to situations where all uncertain inputs are characterized by probability distributions.

4.5.1 Bounding Low Probabilities with Dual-Loop Sampling

First, we consider bounding the failure probability when we loop over epistemic and aleatory uncertainty. This sampling setup implemented in the xLPR code is utilized for all of the scenarios as there are both epistemic and aleatory variables identified in each scenario input set. Looping has the advantage of quantifying the knowledge (epistemic) uncertainty in a probability (risk), but at the cost of losing precision when trying to bound an overall probability including epistemic, aleatory, and sampling uncertainty. When simply trying to bound an overall probability, looping over uncertainties is less efficient than just randomly sampling all inputs without any distinction between aleatory and epistemic.

Recall that in the dual-loop method, n_e epistemic inputs are first sampled in the outer loop. Then, for each sampled epistemic input, aleatory inputs are sampled in the inner loop. While the total number of samples is $n_e \cdot n_a$, the looped approach will generally be less efficient compared to taking a SRS of size $n_e \cdot n_a$. The loss of efficiency when looping occurs because we only have n_e unique epistemic samples, rather than n samples with an SRS. We can think of the looped sample as a ‘cluster sample,’ where we have n_e clusters and n_a samples within each cluster. In survey sampling, the relative efficiency of a cluster sample compared to a simple random sample depends on the amount of similarity within a cluster, often quantified using intra-class correlation (ICC) (Lohr, 2009). In xLPR, the idea of clustering can be mapped back to the sensitivity of the inputs: if most of the variance in the output is explained due to epistemic variables, there will not be much variability in the output due to the aleatory samples (high within-cluster correlation); on the other hand, if most of the variance in the output is explained by the aleatory inputs, there will be a lot of within-cluster variability. In the extreme case, aleatory sampling causes no variance in the response

(within-cluster correlation is 1). The dual-loop method will then result in only n_e unique samples amongst the variables that matter – despite running the code $n_e \cdot n_a$ times.

4.5.1.1 Simulation Example

We will revisit the simple example shown in Section 4.1.2, where we will consider the time to leak T_L in the model studied by (Jyrkama & Pandey, 2016):

$$T_L = T_I + \frac{W}{R} \quad (15)$$

Recall, that given the same distributional assumptions in Section 4.1.2, T_I explained the majority of the variation of the response. We still assume W is fixed and considered two cases specifying the type of uncertainty assigned to the two remaining variables: 1) T_I -aleatory, R -epistemic, 2) T_I -epistemic, R -aleatory. For each case, we consider the overall mean of T_L to be the QoI and assess the variability in the estimate of this QoI under the dual-loop and SRS strategies using simulation. Specifically, we generate 100,000 values of T_L under each sampling strategy and each variable assignment case and compute the mean of these values. This is repeated 1000 times to obtain a distribution of means. The sampling strategy that results in a smaller distributional variance of the mean is more desirable because the estimate has less sampling uncertainty under this strategy. In statistical terms, the estimate is more efficient. For the dual-loop strategy, $n_a = 100$ and $n_e = 1,000$. The distributions of means for each case and each strategy is displayed in Figure 254. In Case 1, the two strategies result in roughly the same variability in the mean. This is because the most important variable is uniquely sampled 100,000 using both the SRS and dual-loop strategies. On the other hand, for Case 2, the dual-loop strategy is much less efficient than SRS since the most important variable is uniquely sampled only 1,000 times (as opposed to 100,000 times) when using a dual-loop.

In general, if the QoI is a summary across both epistemic and aleatory distributions, the dual-loop strategy can only result in a loss of efficiency in the corresponding estimate of the QoI.

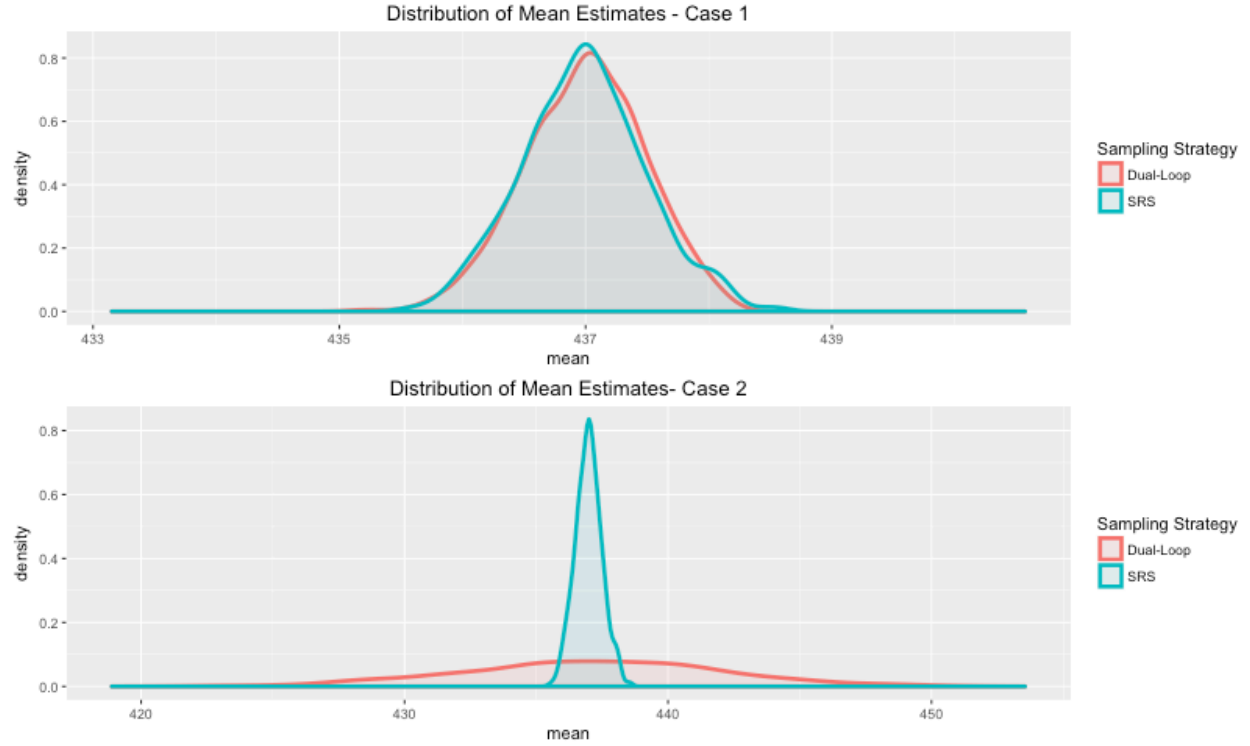


Figure 254: Distribution of means for each case and sampling strategy. The top plot is Case 1 where T_1 is aleatory and R is epistemic. The bottom plot is Case 2 where T_1 is epistemic and R is aleatory.

4.5.2 Bounding Low Probabilities without Dual-Loop Sampling

When the dual-loop is not implemented, giving an upper bound on the probability of an event given zero observed failures is relatively straightforward. Suppose X is a binomial random variable, where X is the number of ruptures in n independent Bernoulli trials. If there are no observed ruptures, an upper bound on the probability of rupture (p) with $100(1-\alpha)\%$ confidence can be found using Equation (16).

$$p_{Upper} = 1 - (\alpha)^{\frac{1}{n}} \quad (16)$$

For example, if 1,000,000 realizations are run and there are no ruptures, one could say the probability of rupture is less than $\approx 3e-6$ with 95% confidence. Solving this equation for n gives the number of samples that are needed to bound a given probability level (i.e., $1e-6$) with $100(1-\alpha)\%$ level of confidence:

$$n = \frac{\log(\alpha)}{\log(1 - p_{Upper})} \quad (17)$$

Figure 255 shows the tradeoff between the number of samples needed without rupture and the given level of confidence for $p_{Upper} = 1e - 6$. As the confidence level increases, the number of samples needed to bound the probability of rupture rapidly increases. For instance, if we wanted to state with 95% confidence that the probability of rupture is less than $1e - 6$ we would need roughly $3e6$ model runs without observing a rupture.

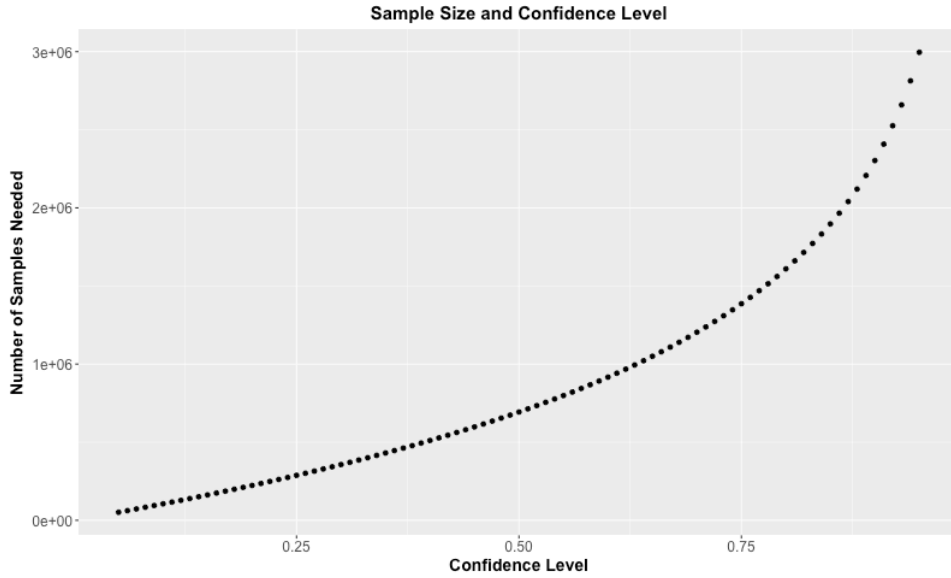


Figure 255: Number of samples needed to bound a probability of $1e-6$ at varying levels of confidence.

4.5.3 Running xLPR Inputs without Dual-Loop Sampling

The xLPR code contains a dual sampling loop by design. This sampling loop cannot, and should not, be removed from the xLPR code. However, changes to the xLPR input set file can produce results that are generated as if they are not created using this dual-loop structure.

The aleatory sampling scheme used within the xLPR code changes the aleatory sample taken for each aleatory realization within a single epistemic realization. The sampling sequence used to generate n_a samples for each epistemic realization is not repeated for the next epistemic realization. Thus, each epistemic realization represents a unique set of aleatory samples.

This characteristic of the aleatory sampling scheme in xLPR can be used to generate a large number of results. If all of the uncertain inputs in the input set are set to aleatory, each aleatory realization will represent a unique sample for all uncertain inputs. Thus, it is possible to sample all uncertain inputs $n_e \cdot n_a$ times. This is the same as sampling the variables as if they are not running in a dual-loop structure.

In order to generate $1e6$ realizations, ten simulations were run with different randomly-selected epistemic and aleatory random seeds. These simulations each had an epistemic sample size of 1000 and an aleatory sample size of 100 with all variables set to aleatory. Again, this allowed each aleatory realization to be a unique sample for all uncertain inputs.

This sampling scheme allows for the generation of a large sample size while also overcoming memory limitations; the xLPR code is only able to run a limited number of epistemic realizations

before the size of the saved results becomes too large for the Windows OS to handle. At the same time, the number of aleatory realizations per epistemic realization is also limited by memory; variable allocation within the aleatory loop cannot support extremely large sample sizes. Running many epistemic realizations with a certain number of aleatory realizations and combining results from sets of simulations with different random seeds allows xLPR analysts to overcome these memory limitations.

Under this sampling structure, there were not any realizations that experienced an occurrence of rupture with LRD. In order to ensure that this result was not due to a low detectable leak rate (set to 1 gpm for all 11 scenarios) these runs were completed with a detectable leak rate of 10 gpm. This change to the detectable leak rate did not produce any realizations with an occurrence of rupture with LRD.

4.5.4 Bounding Low Probability Events Summary

The dual-loop Monte Carlo method is not optimal for bounding low probabilities (across all uncertainties), because the efficiency decreases as the percent of variance explained by epistemic variables increases. However, dual-looping will be appropriate when the goal is to calculate a knowledge uncertainty bound on a probability. As an example, if we wanted to show that at least 95% of epistemic realizations were less than an upper probability bound, looping would be required. Such looping will likely be computationally infeasible for estimating small probabilities, as we would need sufficiently high sample sizes within each epistemic loop to accurately estimate the failure probability. For instance, with $p < 1e-6$, we would need $\gg 1e6$ samples within each epistemic loop depending on the confidence level desired.

In summary, to bound low probabilities across all uncertainties, a non-looped SRS will be more efficient than looping, with a sufficient number of realizations to bound the desired probability at a given level of confidence.

This page intentionally left blank.

5 SCENARIO ANALYSIS SUMMARY

The results of the xLPR Scenario Analysis effort serve as a proof of concept for similar probabilistic studies using the xLPR code. The scenarios analyzed in this report exercised many of the various simulation options and model components comprising xLPR v2.0. The simulation workflow developed for the purposes of this analysis demonstrates the systematic use of the sampling options available for use in the xLPR v2.0.

As expected, increases in epistemic sample size were found to decrease result uncertainty. In all cases, if the goal of a similar study is to estimate the lowest probability outputs (i.e., probability of occurrence of crack ruptures), larger sample sizes than those presented in this report may be desired. The selection of this sample size depends on user requirements for characterizing the uncertainty in their results. The width of sampling confidence intervals can be used as a metric for this selection but, as was shown in the analysis results from this work, increases in sample size may eventually only make small incremental differences in the confidence interval width. In this report, we address choosing a sampling scheme for obtaining a precise best-estimate of the probability of failure across all epistemic and aleatory samples, rather than bounding the probability of failure across epistemic uncertainty. Orders of magnitude larger sample sizes and/or more targeted importance sampling may be necessary for accurately bounding the probability of failure for regulatory decision-making.

The use of Latin Hypercube Sampling (LHS), as opposed to simple random sampling (SRS), consistently proved to provide sampling that ensured stable coverage of the input space and produced the most converged results, as was supported in the discussion in each scenario section in Chapter 3. Importance sampling was shown to have a positive impact on increasing the number of event occurrences; the selection of the importance sampling distribution via the importance sampling quantile impacted the observed efficiency gains with importance sampling. The use of sensitivity analyses, coupled with corresponding support from scatter plots showing the relationship between inputs and outputs, was found to be a useful tool for choosing variables for importance sampling.

The statistical methods described for use in the xLPR Scenario Analysis effort can be used to analyze future xLPR results, including methods of estimation of the mean and quantification of sampling uncertainty. The use of sensitivity analysis methods provides information on the variables that are driving results for each specific input set and helps to support the selection of variables for importance sampling. The calculation of epistemic quantiles on QoIs facilitates bounding epistemic uncertainty in the QoIs. Prediction intervals characterize run-to-run variability in the results as the random seed changes. These tools can be leveraged as post-processing capabilities for more general studies using xLPR results.

This page intentionally left blank.

6 REFERENCES

- Casella, G., & Berger, R. L. (2002). *Statistical Inference* (Vol. 2). Pacific Grove, , CA: Duxbury.
- Davison, A. C., & Hinkley, D. V. (1997). *Bootstrap methods and their application*. Cambridge, New York: Cambridge university press.
- Davison, A. C., Hinkley, D. V., & Young, A. G. (2003). Recent developments in bootstrap methodology. *Statistical Science*, 141-157.
- Efron, B., & Tibshirani, R. J. (1994). *An Introduction to the Bootstrap*. CRC Press.
- Firestone, M., Fenner-Crisp, P., Barry, T., Bennett, D., Chang, S., Callahan, M., . . . Barnes, D. (1997). *Guiding Principles for Monte Carlo Analysis*. Washington: Environmental Protection Agency.
- Hahn, G. J., & Meeker, W. Q. (2011). *Statistical intervals: a guide for practitioners* (Vol. 328). John Wiley & Sons.
- Harrell, F. E. (2016, 12 31). Hmisc: Harrell Miscellaneous.
- Helton, J. C., & Davis, F. J. (2003). Latin Hypercube Sampling and the propagation of uncertainty in analyses of computer systems. *Reliability Engineering and System Safety*, 23-69.
- Helton, J. C., Johnson, J. C., Sallaberry, C. J., & Storlie, C. B. (2006). *Survey of Sampling-Based Methods for Uncertainty and Sensitivity Analysis*. Sandia National Laboratories.
- Hyndman, R. J., & Fan, Y. (1996). Sample quantiles in statistical packages. *American Statistician*, 50, 361–365.
- Iman, R., & Conover, W. (1979). The use of the rank transform in regression. *21*(4), 499-509.
- Jyrkama, M. I., & Pandey, M. D. (2016). On the separation of aleatory and epistemic uncertainties in probabilistic assessments. *Nuclear Engineering and Design*, 303, 68-74.
- Lohr, S. (2009). *Sampling: design and analysis*. Nelson Education.
- Manteufel, R. (2000). Evaluating the convergence of Latin Hypercube Sampling. *41st Structures, Structural Dynamics, and Materials Conference and Exhibit*.
- McKay, M. D., Beckman, R. J., & Conover, W. J. (1979). Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 239-245.

- Neter, J., Kutner, M. H., Nachtsheim, C. J., & Wasserman, W. (1996). *Applied Linear Modeling* (Vol. 4). Chicago: Irwin.
- R Core Team. (2016). R: A Language and Environment for Statistical Computing. Vienna, Austria. Retrieved from <https://www.R-project.org/>
- Saltelli, A., Chan, K., & Scott, M. E. (Eds.). (2000). *Sensitivity Analysis*. New York: Wiley.
- Stein, M. (1987). Large Sample Properties using Latin Hypercube Sampling. *Technometrics*, 239-245.
- Storlie, C. B., Swiler, L. P., Helton, J. C., & Sallabery, C. J. (2009). Implementation and evaluation of nonparametric regression procedures for sensitivity analysis of computationally demanding models. *Reliability Engineering and System Safety*, 94(11), 1735-1963.
- U.S. NRC. (2015). *10 CFR Part 50, Appendix A, General Design Criteria for Nuclear Power Plants*. Washington, D.C.
- xLPR-SDD-FW. (2016). *xLPR Software Design Description for xLPR Framework, Version 3.0*.
- xLPR-SRD-FW. (2016). *xLPR Software Requirements Description for xLPR Framework, Version 4.0*.

7 DISTRIBUTION

1	MS0748	D. Brooks	6233
1	MS0748	A. Clark	6231
1	MS0747	R. Dingreville	6233
1	MS0744	A. Eckert-Gallup	6233
1	MS0829	L. Hund	436
1	MS0830	J. Lewis	436
1	MS0747	P. Mariner	6224
1	MS0829	N. Martin	436
1	MS0748	S. Sanborn	6233
1	MS0899	Technical Library	9536 (electronic copy)

